万卷方法

TELEPHONE SURVEY METHODS
SAMPLING, SELECTION, AND SUPERVISION

电话调查方法

抽样、选择和督导

■保罗·J·拉弗拉卡斯　著
■沈崇麟　译

第2版

重庆大学出版社

责任编辑：邱　慧　方　正
封面设计：黄　河

用什么方法才能最有效地实施和监督电话访谈并减少拒答率？如何进行随机电话拨号？如何选取合适的应答者并获取他们的合作？本书的这一最新修订版详细地讨论了这些问题，为研究者提供了一个思考和实施电话调查的基本指导。

本书涵盖了从抽取电话号码到组织访谈员、监督访谈实施的全部过程，这一版本特别集中讨论了CATI（计算机辅助电话调查）与PAPI（纸笔调查）的比较、混合模式调查、新的电话交流技巧等问题。此外，本书还包括了诸如如何控制样本库、如何区分出适合的应答者以及如何改善访谈员培训等问题。



ISBN 7-5624-3441-7
定价：15.00元

TELEPHONE SURVEY METHODS
SAMPLING, SELECTION, AND SUPERVISION

电话调查方法

抽样、选择和督导

重庆大学出版社

图书在版编目(CIP)数据

电话调查方法:抽样、选择和督导/(美)拉弗拉卡斯(Lavrakas,P J)著;沈崇麟译 —重庆:重庆大学出版社,2005 8
书名原文:Telephone Survey Methods:Sampling,Selection and Supervision
ISBN 7-5624-3441-7

Ⅰ.电 Ⅱ ①拉 . ②沈 .. Ⅲ 电话—调查方法 Ⅳ C31

中国版本图书馆 CIP 数据核字(2005)第073699号

电话调查方法:抽样、选择和督导
(第2版)
保罗 J 拉弗拉卡斯 著
沈崇麟 译
责任编辑 邱 慧 方 正 版式设计 周 晓
责任校对 邹 忌 责任印制 秦 梅
*
重庆大学出版社出版发行
出版人 张鸽盛
社址 重庆市沙坪坝正街174号重庆大学(A区)内
邮编 400030
电话 (023) 65102378 65105781
传真 (023) 65103686 65105565
网址 http //www cqup com cn
邮箱 fxk@cqup com cn (市场营销部)
全国新华书店经销
重庆科情印务有限公司印刷
*
开本 890×1240 1/32 印张 6 625 字数 194千 插页 32开2页
2005年8月第1版 2005年8月第1次印刷
印数 1—4 000
ISBN 7-5624-3441 7 定价 15 00元

作译者简介

保罗·J·拉弗拉卡斯　芝加哥罗要拉大学博士，美国西北大学新闻和统计学教授。1980年以来一直执教于该校，1982年创建西北大学调查实验室，并任实验室主任至今。1968年毕业于密西根州立大学，获社会科学学士学位，并于1975年和1977年在芝加哥罗要拉大学获实验社会心理学硕士学位和应用社会心理学博士学位。

在过去的15年间，由于他在公共政策研究及市民对犯罪的反应和社区反犯罪项目方面取得的成就而在国际学术界声名日隆。他还经常为美国司法部和私营机构在各种有关研究方法论的问题上提供咨询。而在最近6年，他把研究的目光转向媒体如何将调查方法用于新闻事件的调查，特别是政治竞选中的新闻事件的调查。他在这方面的工作也已有所成就，他是《选票和总统选举》(*Polling and Presidential Election Coverage*)(1991)一书的编者之一。他参与过美国中西部舆论调查协会有关美国副总统/总统选举的调查。

沈崇麟　中国社会科学院社会学研究所研究员，中国社会科学院研究生院教授，“第二、三、四批百县市国情调查问卷调查”、“中国城乡居民生活调查”、“中国城乡家庭变迁调查”和“中国城乡社会变迁调查”等研究项目主持人。《社会统计学》、《社会统计方法和技术》、《实用社会科学计算机技术》以及第二、三、四批中国百县市国情调查问卷调查的《调查报告和资料汇编》、《世纪之交的城乡家庭》等译著和专著的译者或作者。

为方法理性鼓与呼*

——万卷方法策划报告暨出版说明

究竟是中国内地学界对于研究方法的漠视，导致研究方法出版物的匮乏？
还是研究方法出版物的匮乏，导致学界没有对方法问题予以应有的重视？
这是一个鸡生蛋蛋生鸡的问题。
作为图书出版的编辑人和策划人，对之多论无益。
但，作为图书出版的编辑人和策划人，我们却非常乐意——为方法理性鼓与呼！

我们乐于，也更善于从出版物的数量和质量的比较中，来了解和表述某一类知识的生产和传播情况，以此作为我们图书出版策划的基础。同样，在万卷方法的策划之前，我们便对美国、中国台湾及中国内地三者，1999—2002年间关于社会科学研究方法的图书出版情况做了一个简单的比较：

美国在社会科学研究方法方面出版物的数量之多，至少足以让我们惊讶！由于不便统计，在此，仅就美国的SAGE出版公司在社会科学研究方法方面的图书出版情况，做一个简单介绍，以窥其一斑。SAGE出版公司4年中出版的社会科学研究方法方面的书籍（包括再版书）便有两百余种，其中，既有一两百万字的大部头工具书，如*Handbook of Qualitative Research*, *Handbook of Research Design & Measurement*；也有10万字左右的口袋书，如仅一个"应用社会科学研究方法系列丛书"（*Applied Social Research Methods Series*）便有49个品种，其中每本书对于案例研究方法、调查研究

* 本文原载《中华读书报》2004年3月31日，第15版。在此，根据需要作者对原文做了少量改动。

方法、网络调查方法等做了分门别类的介绍。其覆盖面之广，研究之深入、具体、操作性强是我们所难以想像的。如果说，我们与美国的差距尚可理解和接受的话，那么，同一时期，中国台湾在这一方面的研究和传播情况也足以令我们汗颜。4 年间中国台湾出版了大量深入研究社会科学研究方法方面的书，其中不但有本土作者的原创作品，也有从英美等西方国家引进版权的相对比较成熟的社会科学研究方法方面的作品。更为可贵的是，其中有几家出版公司已经注意从社会科学研究方法的体系着眼来组织自己的策划和出版，在这方面的出版有了一定的规模，内容的组织也显得比较成熟：如韦伯文化事业出版社的“研究方法系列丛书”、弘智文化事业公司的“应用性社会科学调查研究方法系列丛书”等都是发展比较成熟、规模比较大的专门传播社会科学研究方法的系列丛书。

而中国内地方面，同期虽然也出版了一批研究方法的书籍，但无论是规模还是研究内容的深入丰富程度都无法与美国及中国台湾两地相提并论。从规模看，这期间中国内地出版的社会科学研究方法类的著作也就四五十本，难望美国同行之项背，仅与中国台湾同期出版规模大致相当。从此类出版物内容的深入丰富程度来看，内地出版的社会科学研究方法类出版物主要集中在各学科内部，如教育研究方法、心理学研究方法、社会学研究方法、经济学研究方法、体育科学研究方法等，每本书都面面俱到谈调查、谈案例、谈访谈、谈田野、谈观察等，而很少有对某一种方法进行深入研究的图书，更没有像美国或中国台湾那样形成比较系统的研究社会科学研究方法的系列丛书。在这一领域，这样的图书结构对于应付大

学本科生在研究方法方面的入门性需求（即作为教材）还行，但对于迅速培养一大批拥有科学、理性的研究头脑的学者，特别是对于青年学者，如硕士、博士研究生的成长则是远远不够的。其致命的弱点表现在三个方面：其一，一本书铺得太开而无法细化深入，以至于绝大多数学生虽然上了方法的课，却除了记住了几个名词和对一些方法的理论及应用略知一二之外，至于什么情况下选择什么方法最恰当、具体怎么操作、这种方法可能会有哪方面的不足需要加以处理等研究细节问题仍然处于无理性状态。其二，既然我们的方法建设囿于学科内部，而不能对方法进行纵深的开发，那么，反正学科就那么几个，于是方法书大多是低水平重复。比如教育研究方法的书，现在全国使用做教材的至少也有十多本吧，但你若有心思看的话，一本足矣！其三，出版界对研究方法图书这种淡淡的一笔带过的做法，不仅很难形成一种风气，从意识深处提升起大家对于研究方法的关注，而且更加剧了整个学界对于研究方法的漠视，许多学者只有在著书立说之时才想到似乎从“作品结构的完整性”上看应该谈及方法问题。

社会科学界近年流行两个词语：一曰反思，二曰接轨。所谓反思也即是对于学科的理论基础、学科的结构、学科的历史与未来等问题的全面梳理；所谓接轨也即是希望中国的社会科学研究能够融入世界社科研究的大潮中，与国际上的社会科学研究实现彻底的、清晰的对话。在此，我们以为，无论是反思还是接轨，中国的社会科学界和传播界都必须投入一些精力来关注、研究、教授和传播社会科学研究方法。惟如此，才能在中国形成完善、科学的社会科

学研究方法体系和学科群，也才能使对研究方法关注与理性应用在中国社会科学界深入人心、蔚然成风；惟如此，也才能为中国社会科学研究与国际接轨搭建一个平台。

以上种种，促使我们策划出版了万卷方法系列丛书，献给忠实于人文社会科学研究的人们！

雷少波　崔祝

2004年6月于重庆大学出版社

英文版前言

20世纪80年代中期，在我撰写本书第1版的时候，抽样调查研究方法在电话调查中的应用才刚刚起步。从此之后，随着在研究方法方面许多重要研究的进展和许多研究报告的面世，数量众多的新研究方法不断出现。其中鲍勃·格罗夫斯(Bob Groves)和他的同事比墨、格罗夫斯、莱伯格、马修维茨和舒德曼在1991年 (Biemer, Groves, Lyberg, Mathiowetz, & Sudman, 1991)，格罗夫斯在1989年(Groves, 1989)，格罗夫斯、比墨、莱伯格、马赛、尼可尔斯和沃克斯伯格在1988年(Groves, Biemer, Lyberg, Massey, Nicholls, & Waksberg, 1988)的工作所产生的影响尤为引人注目。他们认为任何严谨的从事抽样调查的研究人员将要或者说必须要考虑抽样调查的计划、执行和结果的解释等各种问题。今天，一个见多识广的决策者，无论是在调查之前的设计过程中，还是在对一个已经完成的调查结果进行解释的时候，都必须对各种对调查精度产生影响的因素，如调查范围误差、无回答误差、测量误差和抽样误差加以细致的考虑。

本书第 2 版较第 1 版有了很大的修改，它反映了（希望如此）我自己在电话调查的计划和执行方面的一些思索。而所有这些都得益于自本书第 1 版问世以来的 8 年中，我在电话调查方面取得的经验，以及自 20 世纪 80 年代中期以来有关总调查误差问题的学习和写作（参见 Fowler, 1993; Groves, 1989）。由于上述诸方面的心得体会，使这一版较之原来的那一版，导向性主题变得更为清楚：运用电话调查的目的在于减少可能的误差源（即偏倚和变差），以

使调查结果达到高质量（即高精确度）。

本书的修订版的目的和本书的原版完全一致：为那些可能考虑使用或正在进行电话调查的人提供一种详尽而实用的方法。我一直都相信，研究人员在考虑采用何种方式搜集调查数据时，在大多数情况下，他们首先应该就是否使用电话调查做出明确的决定，如果拒绝使用它是因为它不适合调查的目的，或无实际实施的可能。正因为如此，只要合适且可行，电话调查总是应该成为各种可供选择的调查方法的首选，因为它给我们提供了更多的机会来系统地控制整个调查过程的质量，从而使我们有更多的机会来减少总调查误差。

不过，这样说并不意味着调查方式的选择必须是非此即彼的。诚如本书所介绍的那样，人们已经越来越多地使用所谓的混合式调查，即在数据搜集过程中同时使用一种以上的调查方法（如电话调查和邮寄调查）。今天，研究者不应当把各种不同的数据搜集方法视为互不相容的，相反，我们应该把它们结合起来使用，以臻扬长避短、彼此补充之效。

尽管有些优秀的教科书已对调查研究方法作了一般的介绍，但其中专门介绍电话调查的并不多见，而像本书这样从实际应用的角度对它进行介绍的更是凤毛麟角。本书对作为搜集数据方法的电话调查，无论是纸笔式电话调查，还是计算机辅助电话调查的具体步骤和规范作了平实而十分详尽的介绍。如果我们想搜集到高质量的数据，那么在电话调查的设计和执行过程中，必须对本书所

介绍的有关电话调查的一般步骤和特定规则予以特别的关注。有许多作电话调查的人似乎并不了解如何建立最有可能确保抽样的有效性和访谈标准化的规范，而这二者正是所有高质量的电话调查所追求的目标。有许多人虽然已经认识到建立诸如这样的规范的重要性，但他们似乎并不愿和/或不能为之而努力。

限于篇幅，我们不可能在本书中介绍电话调查方法的所有方面。本书将主要选择那些不常为其他教科书所详细介绍的部分——被调查人的抽样问题、被调查人的选择问题和调查员的督导问题进行较为深入的介绍。特别要提醒读者注意的是，本书并未涉及用于电话调查的高质量问卷的设计问题。相反本书以大量篇幅介绍如何对调查员的问卷使用进行质量控制。因为高质量的问卷设计，并不能确保数据的高质量，除非我们能辅之以高质量的抽样和高质量的访谈。

与本书第 1 版一样，我仍然假定本书的绝大多数读者都打算计划并实施地方性（如一个城镇、县或州）而非地区性或全国性的电话调查。同时我还假定，绝大多数读者并不拥有那些大型学术单位或私营部门的调查机构才拥有的那些资源。正因为如此，虽然本书用了一定的篇幅来介绍二阶随机拨号（RDD）抽样，但它主要关注的却是地方层次上的抽样库生成和处理方法。在未积累丰富的地方层次上的电话调查的经验之前，我们不应贸然从事地区性或全国性的电话调查。

本书虽然多处提及计算机辅助电话访谈（CATI）与传统的纸

笔式访谈（PAPI）之间的相似和相异之处，但并未对它展开充分的讨论。在过去的十余年间，CATI 有了很大发展，而且还会有更大的发展。虽然如此，它仍然是一种比较新且正在发展中的技术，因而我相信它的使用应限于那些已经在传统的 PAPI 调查中有丰富经验的研究人员。

我想再一次表示对列昂纳德（Leonard）、贝克曼（Bickman）和黛布拉·罗格（Debra Rog）等人的感谢之情，感谢他们给我提供了撰写本书的机会和在本书撰写过程中给我提供的帮助。此外，我也想借此机会对我在西北大学调查实验室以前和现在的同事朱迪思·A·斯凯巴尔（Judith A. Schejbal）、桑德拉·L·保玛（Sandra L. Bauman）、苏珊·M·哈特内特（Susan M. Hartnett）、J·索菲·玻克南（J. Sophie Buchanan）和丹·M·迈克尔（Dan M. Merkle）的工作能力表示敬佩，并感谢他们在研究工作中给我的支持。

我还想借此机会表达我对鲍勃·格罗夫斯（Bob Groves）和西摩·苏德曼（Seymour Sudman）等人的敬佩，我与前者虽然从未谋面，但他对我在高质量的调查方法方面的思想影响极深。后者不仅对抽样调查方法的贡献巨大，而且他在这方面表现出来的高度专业水准和亲力亲为的风范，堪称几代调查研究工作者的楷模。此外，我将永远铭记和感谢列恩·贝克汉姆（Len Bickman）、鲍勃·包罗克（Bob Boruch）、玛里琳·布鲁尔（Marilyn Brewer）、玛高·戈登（Margo Gordon）、彼得·米勒（Peter Miller）、埃米尔·波萨发克（Emil Posavac）、丹尼斯·罗森博姆（Dennis Rosenbaum）、富兰

克·斯莱马克尔（Frank Slaymaker）、韦斯·斯科甘（Wes Skogan）和迈克·特罗格特（Mike Traugott），感谢他们给我的鼓励和实践机会。

最后，我想对我的妻子芭芭拉·J·拉弗拉卡斯（Barbara J. Lavrakas）和儿子尼古拉斯·J·拉弗拉卡斯（Nikolas J. Lavrakas）说一声谢谢，感谢他们对我献身我所感兴趣的事业所表示的理解。

保罗·J·拉弗拉卡斯

目　录



导　论

Introduction

总调查误差
电话调查面面观
电话调查在美国
电话调查的基本步骤
计算机辅助电话访谈(CATI)
本书的内容和结构
有关调查方法的其他知识来源简介
练　习

人们对量化和以某种可测量(数字)的方式描述某些事物的方法感兴趣可谓由来已久。数量的内在意义不仅仅限于它们的数字符号所表示的量。虽然在历史早期,一般老百姓很少需要对数量问题有广泛的认识,但在当今社会生活中却要求人们对数量问题有一定的了解。即使是那些认为自己是没有数量概念的人也不得不承认,他们日常生活中的很多决策都是基于某种数量上的考虑。

长时间以来,人们已经认识到有效的决策并非总是需要确切的计数或测量,而这种认识正是抽样理论发展的基础。在对事物的认识和决策过程中,人们很早就已经懂得使用抽样,或只考虑较大集合中所有可能单元、人口、地方、事物等总体的一个子集等方法(例如,这些方法就是所谓典型概括法的基础)。然而,用于系统抽样的正规方法的产生则距今不远。正规抽样方法的产生,才使那些我们叫做抽样调查方法(survey methods)的各种调查技术应运而生。

抽样调查方法是一组调查技术的统称。这些方法最基本的目的是为我们提供我们感兴趣的某些变量的普遍程度(即数量)的精确估计值(即一些比较稳定且方差相对较小的量度)。例如在已经登记的选民中,赞成某一候选人的百分比是多少;所有的住户中,拥有摄像机的住户的百分比是多少;在所有从某一组装线上下线的小轿车中,有结构性缺陷的百分比是多少。在这些事例中,为了能对谁将赢得选举、摄像机的市场大小或某一特定汽车组装厂的工艺质量等问题做出精确的判断,我们不必无一遗漏地去调查所有可能的选民、住户或小轿车,以搜集作为我们判断依据的各种资料。

只要抽样调查的方法运用得当,无论调查对象是什么,我们都可得到具有代表性的非常精确的估计值。这些估计值的精确度可达到误差忽略不计的程度,至少从有效的决策角度看情况是这样的。不过有效的(极可能是精确的)抽样调查方法涉及很多新的且正在发展过程中的知识,正因为如此,它们常常被误用和误释。此外,我们还必须了解,抽样误差仅仅是诸多可能的误差源中的一种。

总调查误差

格罗夫斯(Groves,1989)在他著名的有关抽样调查研究中各种可能的误差源的论文中明确指出,在抽样误差之外,从事抽样调查的严谨的研究人员,必须考虑选择那些对范围误差(coverage error)、无回答误差(nonresponse error)和测量误差(measurement error)的可能影响已加以控制和/或测量的方法。正因为如此,我们始终认为那种只考虑抽样误差的思想是"不全面和不能令人满意的"(Groves,1989,p.13)。因为总调查误差并非只是由抽样误差,而是由包括抽样误差在内的所有的不精确源和偏倚源一起形成的(参见 Fowler,1993;Fowler & Mangione,1990),不仅如此,我们还必须在计划、执行和解释一个抽样调查时,对各种可能的误差源给予分别的考虑。

抽样误差

在对抽样调查得到的结论的效度和实用性进行评估时,我们经常会遇到那样一些人,他们所考虑的问题主要(甚至只)是抽样误差的大小。抽样误差是我们所测量的事物的异质性、样本和总体大小的函数。抽样调查中之所以会出现这样一种误差,是因为与普查不同,它观测的并非总体中的所有单元,而只是其中的一个样本。例如,在用一个样本容量为 500 的随机调查,对一个较大总体(譬如 >10 000)中的某一分布约为 50/50 的现象进行观测时,我们可以确定在 95% 的置信水平上,抽样误差约为 ±4%。不过,我们并不能就此确定调查对这一现象测量的精确性一定在 4 个百分点之内,因为在抽样误差之外,还存在着许多可能的偏倚源和误差源。我们常常把这些其他的误差源统称为非抽样误差。

范围误差

抽样调查得到的结论可以推广到样本赖以产生的总体,但在对结论的精确性做出判断之前,我们必须考虑这样一个问题——

总体中的所有单元是否都有被抽入样本的某种可能(概率)? 例如,在所有的一般公众中进行电话调查时,那些没有电话的住户和无家可归的人被抽入样本的概率都为零。因此,在我们试图将电话调查的结论推广到整个公众时,有可能会受到范围误差的影响。只要我们所研究的现象在一定程度上与覆盖或未覆盖范围有关,调查结论的精度就会因此而降低。在电话调查中,因为那些无法通过电话联系的公民,作为一个群体,与总体中其他有电话的群体相比,他们在收入或其他与收入有关的社会经济地位上一般都处于比较低的层次,所以在电话调查中从被调查人那里得到的有关收入或其他与收入有关的行为的结论可能会高于整个总体。

无回答误差

很少有抽样调查的回答率能达到100%,这就是说,大多数调查或多或少都会抽到一些搜集不到数据(其原因是多种多样的,如拒绝回答、假期等等,本书以后的各个章节都会涉及这一问题的讨论)的单元(如人或住户)。如果那些被抽到但没有搜集到资料的单元系统地不同于那些被抽到且搜集到数据的单元,那么调查结论便会在一定程度上存在无回答误差的影响。例如,青年男性不成比例地拒绝参加一项调查和/或因为他们很少在调查员给他们打电话时在家,那么在整个总体中,任何涉及与年龄和性别有关的量度(如泡酒吧或从事身体接触性的运动)的精度,就会因为无回答误差的影响而比较低(即产生偏倚)。如果调查中搜集到的数据与某一特定类型的被调查人是否给出了回答无关,那么因为无回答而造成的误差(偏倚)便不复存在。

测量误差

并非所有记录在问卷上的数据都精确地测量了我们所感兴趣的现象。这些不精确性可能是由与问卷、调查员、被调查人和/或数据搜集方式有关的误差引起的(参见 Biemer, Groves, Lyberg, Mathiowetz, & Sudman, 1991)。例如,问卷中某一调查问题的措辞不当,某一组或全部问题的顺序编排不当都可能导致被调查人误报(偏倚)答案;调查员的行为方式可能使被调查人的回答产生偏倚;

被调查人可能不愿或不能精确地回答某一问题和/或搜集数据使用的方式(如个别访谈和电话访谈)也可能导致测量误差的产生。

调查费用

正如格罗夫斯(Groves,1989)所说,降低和/或测量各种形式的调查误差的影响的努力都意味着调查费用的增加。读者必须了解旨在降低可能误差的方法与旨在测量它们的可能影响的方法之间存在的基本差别。采用那些可以消除(或明显减少)可能误差源的各种方法,费用也许太高,但是采用一种可以测量误差的近似大小而又使我们可以对调查结果进行必要的解释的方法,其费用则可能是我们能承受的(水平较高的读者和调查工作者,可以研读或再次研读格罗夫斯于1989年(Groves,1989)发表的极具挑战性,但极为出色的著作《调查误差和调查费用》(*Survey Errors and Survey Costs*))。

对于那些调查业中的新手来讲,这些有关调查费用和各种调查误差的考虑不免令人望而生畏,甚至感到重压在身。面对有关调查有效性的种种质疑,有些人可能会因此甩手不干,并对整个调查业存在的价值产生怀疑。但是持有这种观点的人也许忽略了这样一个事实:那些经验丰富的调查工作者,不仅关注调查面临的各种问题,而且规范又高度精确地完成了各种调查。

作为一本电话调查导论,本书只是对电话调查应用中出现的一些问题做了一些基本的介绍。我们这样做的目的并不是想降低高质量的调查应有的声誉,更不是想阻止任何人去做一个优秀的调查。相反,我们只是想提醒读者,在执行一个电话调查时遇到的许多问题,就他们调查的目的而言,本书所提供的知识对于这些问题的解决已经有"足够的精确性"。

我想在这里告诉那些初学者:计划、执行和解释一个调查,最终是可以做到比较精确的。这一过程不仅有它自己的特定方法,而且还需要花费相当多的时间,然而我们为此付出的努力是完全值得的。

电话调查面面观

电话调查方法只是在近25年中才得到迅猛发展,这一点也许会令某些人感到震惊。在此之前,美国住户拥有电话的比例还太低,因此我们还没有正当的理由将电话作为一个有效的抽样工具。一旦美国住户拥有电话的比例超过90%(这一情况发生在20世纪70年代),用电话调查进行访谈可涉及的范围在理论上与个别(面对面)访谈同样全面。不言而喻,电话调查中使用的一系列方法,是在20世纪70年代后期和面对面调查一起发展起来的。

在20世纪80年代中期,电话调查开始变得很普遍,而且成为许多场合中各种调查方法的首选。当一种在公共和私营部门被人们用来帮助他们提高决策效率的搜集信息的有效方法为人们所普遍认同时,这一现象本身就具有方法论上的意义。事实上,人们在市场调查中花费的资金,超过了舆论调查和学术性调查二者之和。

为什么作为一种可为某一感兴趣的主题提供精确测量数据的工具的电话调查会变得如此重要?概括地讲,这是因为在许多情况下,它所具备的优点远远超过了它的缺点(作为对下面进行的讨论的补充,有兴趣的读者可参见弗瑞(Frey,1989,p.76)有关各种调查方式优缺点比较的论述)。

虽然许多人不承认或者没有认识到,与那些早期的调查法相比,迄今为止电话调查最大的优点在于它给我们提供了质量控制的机会。这一机会包括抽样、被调查人的选择、问卷题项的提问、计算机辅助电话访谈(CATI)和数据录入。如果我们没有发现其他重要的足以令我们放弃使用电话调查的原因,那么电话调查所具备的控制调查质量的优点,便足以使它成为调查方式的首选。

本书强调对数据搜集过程进行控制和监视的重要性,因为这是取得高质量数据,进而得到精确估计值的保证。没有其他任何调查方式可以给我们提供这种进行质量控制的机会。当我们能恰当地组织好一次电话访谈时,调查数据便会非常接近无偏倚的标准化水平,而这一目标正是所有高质量的调查所追求的(参见 de

Leeuw & van der Zouwen,1988;Fowler & Mangione,1990)。

在科学的调查方法从个别访谈向电话访谈转变的初期,许多人担心电话访谈搜集到的数据质量低于(偏倚更大和/或方差更大)个别访谈。然而过去10年进行的研究证明,这两种方式搜集到的数据在质量上并不存在什么恒定不变的差别,如果说有什么差别存在,似乎也在随时间的推移而逐渐变小(参见 de Leeuw & van der Zouwen, 1988; Groves, 1989)。例如,特图克尔(Tucker, 1983)发现,电话调查中调查员之间存在的变差略低于个别访谈。格罗夫斯(1989,p. 510)引用了某些研究成果证明,被调查人在回答某些"胁迫性"问题时,如能通过电话,而非直接面对调查员本人的话,那么他们的答案受社会愿望影响的可能(即回答人选择正面倾向的答案的可能)就比较小。另一方面,也存在这样一种倾向,电话调查中被调查者在回答开放式问题时,给出的答案往往都比较简短(参见 Groves,1989,p. 512)。

电话调查的第二个优点是成本效益较高,电话调查的成本效益远远高于个别访谈。例如,电话调查除不需要差旅费之外,格罗夫斯(Groves,1989,p. 512)估计,电话访谈所需的访谈时间,比同样题项的个别访谈少10% ~20%。尽管电话调查的费用较邮寄式调查的费用要高,但它可以降低总调查误差这一优点要远远胜于费用较高的不足。

电话调查的第三个优点是速度快,因为电话调查的数据搜集和处理可以同时进行。我们可以在一个星期或不到一个星期的时间内,通过电话调查搜集到个别访谈可能要用一个月甚至更长的时间才能搜集到的数据。而邮寄式调查所需的时间可能更长,因为为了提高邮寄式调查的回答率,我们必须进行追踪式的邮寄调查,这样的调查往往需要更多的时间。例如10个有经验的电话调查员,4小时一班,在3天之内(其中包括一定数量的重拨)便可完成含20个题项的问卷400 ~500份。假如某一位市长在星期一提出一个任务,希望在周末之前取得某些决策需要的信息(如市民究竟是怎样不满意警察局的工作,她是不是应该解除警察专员的职务),一个优秀的电话调查机构不仅可在成年居民中完成有关这一问题的电话调查,并且还可以在截止日期之前,将调查结果告诉市

长。而邮寄式调查和个别访谈就做不到。

电话调查(即使是组织执行得很好的电话调查)的主要缺点是,受访谈内容的复杂性和时间长度的限制比较大。与面对面的访谈具有较大的能动性不同,电话调查的调查时间,尤其是对老年人调查的时间不应过长,一般不能超过 20 ~ 30 分钟。与电话调查不同,人们似乎比较喜欢与个别访谈的调查员交谈,即使访谈时间超过了 30 ~ 40 分钟,甚至更长,人们也不那么厌烦。邮寄式问卷调查似乎也没有招被调查人讨厌的缺点,如果需要,被调查人可以利用空余的时间,分几次填写问卷。同样,那些需要被调查人看或读的比较复杂的问题,至今仍然无法用电话调查来完成,尽管随着可视电话技术的出现(1992 年起开始向公众出售),这些局限有可能逐步减少。

另一个来自传统调查方法对电话调查的担心是,有些人至今仍然认为,电话调查比较容易产生范围误差和无回答误差。诚如前言,并非每一个人都生活在安装了电话的住户中,而即使是生活在有电话住户的人,在是否愿意/能够给他们打电话并进行电话调查这一问题上不同人口群体(例如老年人口群体)的态度也不尽相同。这就使得电话在联系一般总体中的某些部分的人口时处于不利的地位。不过有几种搜集数据的方法,使我们可以通过电话调查搜集到那些没有电话的人的数据(参见 Frey,1989,p. 46)。这些方法就是所谓的混合式调查,它使我们得以与那些没有电话的人进行联系。我们将在下面的有关章节中介绍这些方法。马克兰和沃克斯伯格(Maklan & Waksberg,1988)对此却持有不同的观点。他们认为一个组织完成得很好的随机拨号(RDD)电话调查,就生活在有电话住户内的人而言,在计数和访问各种类型的人口群体时,所产生的范围误差至少不会高于政府组织进行的大规模调查(例如,统计局每年进行一次的现有人口普查)。对人口中那些有听觉问题的人(特别是老年人),电话调查在无回答误差上仍然存在问题,尽管随着装有助听器的电话通讯设备的推广应用,电话调查中存在的这一问题可能会有所好转。

自本书第 1 版于 20 世纪 80 年代中期问世以来,多框架或混合式调查的使用有了显著的发展,从而促使有关这方面的研究取得

了长足的进步(参见 Dillman & Tarnai,1988,1991;Frey,1989;Groves,1989)。这些调查是使用一种以上的抽样方法和/或使用一种以上搜集数据的方法的调查。例如,一个在某市(伊利诺斯州莫顿格罗夫市)郊区的2 000多个住户中进行的有关学校注册情况的调查中,我们先采用邮寄问卷调查,接着再作电话调查,在遇到需要时,再对那些采用其他调查方法没有得到回答的调查对象作个别访谈,最后使回答率达到了92%(参见 Lavrakas,1990)。

我们不应当这样看待电话调查和其他调查方法之间的关系,仿佛它们之间存在一种彼此竞争的关系。实际上,它们之间的关系不仅不是相互冲突的,而且是可以互为补充的。例如我们可以通过电话来说服邮寄式调查的被调查人接受调查,或寻找一般人口中的稀少群体(如失明的成年人),然后再对他们进行个别访谈(参见 Frey,1989,pp. 24 ~ 25)。与把调查方式的选择看做一种非此即彼的决策不同,现代调查研究者所应寻找的方法是那些可以使不同的调查方式创造性地结合在一起的方法。诚如迪尔曼和塔迈(Dillman & Tarnai,1988)所言:

每一种调查(方式)都有某种难以避免的局限性。……如果不能克服,那么将几种(方式)结合在一起使用,就可以减少单个(方式)存在的某种局限性。混合式调查设计,为我们提供了在扬每一种(方式)之长的同时,避每一种(方式)之短的可能(pp. 509 ~ 510)。

因此从根本上讲,筹划一个调查所面临的挑战是,考虑如何在项目实施条件允许的前提下,通过使用不同的抽样和数据搜集的方式,尽可能地降低各种形式的误差。

电话调查在美国

有两组与电话现象有关的因素,它们对电话调查能否成功(即有效)进行有着直接的影响。第一组因素是“硬件”,它在物质上使我们可以通过电话与被调查人进行联系。第二组因素是社会因素,它们与人们通过电话进行联系的语言的和非语言的行为有关。

此外,随着通讯技术的发展变化和电话推销及电话调查不断地走向公众,我们有可能看到与电话有关的行为的社会规范,会因为这些新的技术和人类的本性及癖好的相互影响而持续不断地发生变化。

我们在这一章所要介绍的内容,只是物质和社会因素中的某些部分。有兴趣的读者,可以参阅德·梭拉珀尔的著作《电话的社会影响》(de Sola Pool,1977);布鲁克的著作《电话:第一个百年》(Brook,1976)。弗瑞(Frey,1989)的著作则对电话现象作了更为详尽的介绍,因为他对电话调查问题作了专门的阐述,而格罗夫斯等人的著作(Groves 等,1988)则论述了人们在电话调查的研究中所取得的进展。

物质因素

电话线路相互连通

20 世纪 40 年代中期,美国的每一个住宅和营业场所的电话,至少是非直接地通过贝尔网(Bell network)相互连接起来的(Pierce,1977)。此前,一个专门的电话调查机构想要通过电话与美国所有电话住户进行联系是不可能的,因为各个独立的电话公司之间的电话线路还没有完全连通。因此,那时用电话号码进行全国性的抽样是行不通的(可能产生范围误差)。随着这些障碍的消除,使美国的电话调查在理论上成为可能,即我们有可能通过电话与任何有电话的住户进行联系,不论调查员在什么地方打电话或给什么地方打电话。

住户电话线饱和度(Saturation of Telephone Lines in Households)

尽管美国在 20 世纪 40 年代已经消除了电话线路不能彼此连通的障碍,但在 50 年代,美国住户电话的普及程度就抽取一个全美人口的有效(有代表性)样本这一目标而言,仍然太低。虽然在美国,这一方面的状况在过去的 35 年间已经有了很大的改观,但这一问题仍然是世界上的大多数国家至今无法有效地使用电话调查的主要障碍。

1960 年,约 80% 的美国住户装有电话,1965 年这一数字上升

到85%,而到了1970年,这一数字已经超过了90%(美国国家统计局,1984)。在20世纪80年代早期,只有瑞典电话与人的比值大于美国(美国电话电报公司,1982)。虽然西北大学电话调查实验室(NUSL)于20世纪90年代初进行的几次随机拨号调查得到的数据显示,住宅电话与非住宅电话已接近2:1,而美国电话电报公司的统计数据显示,在20世纪80年代,住宅电话仅占美国所有电话75%左右。

而最近的统计数字显示,美国至少有一条电话线(即有一个单独的电话号码)的住户在全部住户中所占的比例已经达到了95%。但因为无家可归的问题存在,全美人口中可通过电话联系的人口比例,可能会略微低于这一数字。美国电话电报公司估计1982年有电话的住户的比例为97%,联邦通信委员会报告的1981年的数字为96%(美国国家统计局,1984),而弗瑞(Frey,1989)估计的数字为97%。托姆博瑞和马赛(Thornberry & Massey ,1988)根据全美卫生统计中心每年进行一次的大规模的(40 000户)个人访谈资料,也对这一数字作了估计:在20世纪80年代中期,约93%的美国人生活在有一部电话的住户中。

诚如大家所知,那些没有电话的住户人口学特征并不能在其余有电话的住户中得到反映。没有电话的那些住户通常有这样一些特征:收入和教育程度都非常低,生活在农村,年龄比较小,是少数民族或少数种族。例如,在20世纪80年代中期,所有黑人中的1/6(16%)生活在没有一部电话的住户中,而所有家庭收入在15 000美元年龄在55岁或以上的老年人(老年贫穷人口)中,户内有电话的比例高达97%(参见Thornbeny & Massey,1988,pp. 27 ~ 41)。我们提醒读者注意的一个重要问题是,一些研究表明,在许多行为和态度量度上(除那些关系到收入的量度以外),无电话住户作为一个群体,与人口中的其余群体并无很大的差别(参见Groves & Kahn,1979;Thornberry & Massey,1988),不过这些研究并不包括无家可归人口的数据。

任何计划进行电话调查的人不仅应当对这些问题有所了解,而且还应该对调查结论中由于未覆盖人口问题的影响而可能产生的误差做出明确的判断。对那些在某一地区内进行电话调查的人

来讲,重要的问题是必须了解某一特定地区户内没有任何一条电话线的住户在总住户中所占的比例,这一数字在美国的不同地区是不相同的。联邦政府的有关统计数据显示,在20世纪90年代,有电话住户的比例高于85%而低于90%的州有10个:阿拉巴马(Alabama)、阿拉斯加(Alaska)、阿肯色(Arkansas)、肯塔基(Kentucky)、路易斯安那(Louisiana)、密西西比(Mississippi)、新墨西哥(New Mexico)、俄克拉荷马(Oklahoma)、得克萨斯(Texas)和西弗吉尼亚(West Virginia)(国会信息服务中心,1990)。此外,市内的低收入地区住户的电话拥有率一般都比较低。任何计划进行地方性或地区性一般人口电话调查的人,事先都必须了解抽样区域内电话的普及程度。此后应考虑的最为重要的问题是调查主题的具体性质。在弄清楚这一问题之后,还必须对各种可供选择的调查方式,如电话调查和个别访谈的有效性进行比较,并考虑某种混合调查方式(即电话调查和个人访谈)的可行性。只有在完成这一系列工作之后,才可以做出有充分根据的决策。

美国住户电话线进一步饱和的前景似乎不那么确定。在1984年美国电话电报公司的经营业务被强制性剥离之前,人们担心市内电话设施的费用将会明显上涨,而有电话设施的住户比例将会因此而下降,但是这种情况在20世纪90年代初并没有发生。虽然如此,电话调查的从业人员仍然有必要继续关注住宅电话市场的进一步发展变化。此外,高级研究人员应当考虑用统计修正(加权)来抵消未覆盖误差(noncoverage error)对调查估计带来的可能的影响(参见 Massey & Botman,1988)。

多线路住户

另一个对电话调查有影响的物质因素是多线路住户问题。这一问题对电话抽样产生的影响比没有电话线的住户造成的影响稍微小一些。在采用随机数字拨号(RDD)法抽取住户时,那些有一个以上电话号码的住户被抽取的概率,将会大于那些普通的只有一个电话号码的家庭(注意,这一问题并不是指同一个号码有多个分机。与这一问题有关的住户是那些有两个或两个以上电话号码的,如家庭办公专线、儿童专线等)。

格罗夫斯和卡汉(Groves & Kahn,1979)在全国性的拥有电话

设备的住户样本中发现，拥有一个以上电话号码的住户的比例约为5%；这些户主中，绝大多数都说自己有两个不同的电话号码。在NUSL于1992年进行的全国性调查中，约有14%的被调查户主报告他们拥有一条以上的电话线路。在不同地区，这些数字存在着差异。例如NUSL在芝加哥地区进行年度调查中发现，约有1/5(18%)的住户拥有一个以上的电话号码。在这里，统计修正同样也(且应当)用于纠正因为多线路问题引起的抽样的不等概性。

其他地方性和地区性变差

电话公司的服务站点遍布整个美国，各个服务站点的服务设施和服务政策存在一定差别，而这些差别可能对电话调查的计划和执行产生影响。早在20世纪80年代初，就已经有60家以上的电话公司开设了市内电话服务站点(美国国家统计局，1984)，读者大可不必为此而感到惊讶。例如，各家电话公司都使用不同的程序来开设后缀号码(即电话号码的最后4位数)工作库。不同的市内电话公司，在与电话调查研究者进行配合所持的态度上也存在很大差异。还有，不同的电话公司处理空号的方式也各不相同，这一问题对采用随机数码拨号法的电话调查尤其重要。所有这一切都意味着，那些在某一地区电话调查的抽样中适用的经验，未必适用于另一地区的抽样。

新的电信技术

因为电话行业在不断地开发出新的硬件和软件产品的同时，也在不断推出各种新的服务项目(参见Andrews，1992)，所以电话调查从业者必须随时提醒自己要不断更新调查方法，以适应这些不断进步的技术，从而保证抽样和调查的高质量。

电话应答机的问世和相关服务设施的推出，使电话调查的重呼次数大大高于这项技术产生之前。例如，在20世纪80年代初，在一般公众中进行的电话调查，通常第一次呼叫就可完成全部访谈量的一半(参见Lavrakas，1987)，而现在这一比例已降至1/3左右，造成这一现象的部分原因与应答机的使用和更为忙碌的生活方式有关(参见第3章)。尽管某些感觉告诉我们，在那些使用应答机的人中，绝大多数(至少在20世纪90年代初)主要是因为不

想错过应该接听的来电,而不是想对来电进行筛选,以免接听不想接听的来电。

来电显示是另外一项在20世纪90年代初某些电话公司销售很火的技术。凭借这一技术,人们可以看到来电号码,并由此知道电话来自何方(参见Frey,1989,p. 247)。如果公众对这一技术持普遍欢迎的态度,且大多数可能的被调查人都拒绝接听他们不知道电话号码的来电(即某一电话调查小组的来电),便会给电话调查造成难以克服的困难。而有关这一技术的合法性还存在许多不确定性——有些人认为这一技术的使用侵犯了呼叫者的隐私权——因此这一技术的应用前景还是难以确定的。

转呼、车载电话和其他袖珍电话(例如一种叫做"迪克·特蕾西"(Dick Tracy)的新式的戴在手腕上的电话,在本书第2版撰写期间正处在早期的发展阶段),同样也给公众电话调查的概率抽样带来了麻烦。这些情况都是电话调查工作者在真正实施电话调查之前所必须了解和计划周详的。可以设想,在不久的将来,每个人都可能会有他/她自己的(诸如美国电话电报公司"700"那样的)电话号码,那时的电话调查工作者用它来进行抽样,抽取的将是个人,而不是住户(参见Ramirez,1992)。

也许在今后的10年中,电话调查者所要面临的最大挑战是可视电话的应用,美国电话电报公司已经在1992年便开始以每台1 500美元的价格销售这种产品。同年底,该公司产品的一种仿制品,便以相当于它的一半价格面市。这一技术的普及速度取决于它为公众和经营部门接受的程度,从目前的情况看,它的普及可能还要很多年。而在它真正普及之前,可能会出现一部分被调查人已经采用了这一技术,而另一部分却还没有采用的情况。这势必会给电话调查带来一系列新的问题。如在遇到那些除非在电话中可以看到调查员,否则拒绝在电话中进行访谈的被调查人时,调查员应该怎么办呢?对调查员的着装应该有什么要求呢?调查员应该装备什么样的可视设备呢?与非可视的电话调查相比,可视电话调查会对测量误差产生什么样的影响呢?可视电话调查究竟能给电话调查带来什么样的好处呢?这些是在(或者如果)可视电话技术流行过程中,电话调查者所遇到的许多问题中首先应该仔细

研究并回答的问题。

社会因素

由于我们在理论上存在可以通过电话与所有美国住户进行联系的可能性，因此有必要对电话调查运行的心理和社会环境进行一番考察。诚如弗瑞(Frey,1989)所言，存在着某种有利于电话访谈的与电话有关的传统行为规范。不论所有生活在美国的人是否都成长于一个有电话的家庭，由于自第二次世界大战以来，人们受家庭电话的影响如此之深，致使这些规范几乎支配着整个美国文化。

来电必答

在人们的行为规范中，在我看来是非常重要的而在有些人看来是理所当然的一条是来电必答。这就是说，在美国，绝大多数情况下，只要电话铃声一响，就会有人接听电话，除非当时家中没人。但接听电话的快慢，不同的家庭和不同的个人却存在很大差别。对于这一差别，人们同样也会认为是理所当然的。不管怎样，实际经验告诉我们，如果一个电话被接听了，一半左右的电话会在铃声响了三次或四次后被接听，而90%的电话则在铃声响了七到八次内被接听。因此在电话调查中从来都不存在可能的被调查人会不接听来电这样一个问题。

然而随着新技术的出现，已经有迹象表明，公众中那些用应答机筛选掉他们不愿接听的来电的人数比例有所上升，但上升的幅度不是很大。不仅如此，正如我们在上面所说，来电显示技术的面世，就是为了满足这方面的需要。从某种意义上讲，相当一部分公众将会习惯性地拒绝接听某些来电，电话调查中的无回答问题，以及由此而产生的无回答误差可能会变得严重起来。对这方面有兴趣的读者可参阅格罗夫斯(Groves,1989)著作中全面阐述调查中无回答产生的原因的那一章(pp. 185 ~ 238)。

有礼貌地终止电话交谈

还存在第二种不太明显的有利于电话调查的传统行为规范。正如弗瑞(Frey,1989)所观察到的那样，在电话交谈中，决定交谈时

间长短的是呼叫者,即打电话的人而不是接电话的人。这就是说,呼叫者之所以打电话,总是出于某种目的,而某种礼仪规范告诉人们,接听者应在呼叫者的目的实现之后再挂断电话。显然,并非所有人在接听电话的时候都执行了这一礼仪规范,因为要是那样的话,电话调查员就不会遇到拒绝接听,更不会遇到"不置一词就挂断电话"这样的问题。不过类似后一类问题,确实不是经常发生的。

然而在电话调查中出现的虽然为数不多但却在稳步增长的调查无回答问题说明这一传统的礼仪规范正在趋弱。有鉴于此,我们似乎真应该"感谢"电话推销商,"感谢"他们"驱使"公众学会了"只会说不"。遗憾的是,许多人不知道如何区别合法的电话调查和电话推销。更有甚者,有些电话推销员还试图使用假的电话调查,引接听者上钩。这种做法有时形同"骗销"——就是以调查之名,行推销之实。

真实性判断

另外一个构成电话调查有效性的因素是调查员—被调查人之间的互动具有的真实性。有一种对电话调查很天真的批评,认为被调查人在接受电话调查时经常会说谎,他们因此而断定电话调查得到的数据是无效的。我自己(参见 Lavrakas & Maier,1979;Maier & Lavrakas,1976)和其他人(如 Ekman & Friesen,1974;Ekman & Friesen,1976;Maier,1966;Maier & Thurber,1968)的研究都表明,在听取另一个人的声音和交谈的内容时,聆听者(即调查员)一般都有能力精确地判断她或他谈话的真实性。实际上,这一类研究也表明,个别访谈也许更容易受欺骗,因为以视觉线索为依据,更容易对他人谈话的真实性做出错误的初步判断。

此外,逻辑学的知识也告诉我们,很少会有人愿意为故意欺骗一个电话调查员而浪费自己的时间,尽管某些被误导和非专业的新闻记者一直在惊呼,公众政治性的民意调查就是这样做的。具有这种倾向的人,他们更可能采取的做法是直接拒绝接受访谈。所有已经考虑到的事情告诉我们,在大多数情况下,大多数被调查人在接受电话调查时是愿意提供相当可靠的信息的假定是可信的。

其他社会因素

如同上述,一个社会的人口和行为方式是会发生变化的,而这些变化则可能对电话调查的实践产生影响。例如,因为美国人口中相当大一部分人口是老年人口,因此电话调查中那些与老龄问题相关的问题(听力的丧失、社会脱离等)便会增多。如果其他诸如害怕犯罪和担心隐私受到侵犯这样的社会因素有所增长,那么我们可以设想电话调查的麻烦还会进一步增加。而这些趋势是电话调查研究者所必须考虑的,并且必须找到解决的办法。

小 结

尽管近年来电话调查中的无回答问题有所增长,但到 20 世纪 90 年代初为止,美国社会无论在物质因素还是在社会因素上,都不存在那种会使旨在搜集一般人口或特殊人口(如企业经营者)中有代表性的数据的电话调查方法自动失效的不可克服的障碍。这并不是说电话调查总是各种调查方式的首选,而是说我们应该为电话调查正名,无论从抽样的角度还是从提问的角度看,电话调查都不亚于个别访谈。

电话调查的基本步骤

我们要向任何计划进行电话调查的人强烈推荐的是,应该为每一个电话调查制定一个详尽的操作管理计划。这个计划必须对调查应完成的所有任务做出具体安排,并确定完成每一任务的相应人员(参见 Frey,1989;Lyberg,1988)。下面所介绍的便是进行一个高质量的纸笔访谈(以后将简称为 PAPI,即不借助计算机辅助电话访谈(CATI))式电话调查的基本操作步骤:

1. 确定抽样方案,其内容应包括确定抽样单位抽取的抽样框和在抽样单位中选择被调查人的方法;
2. 确定用于从抽样框中产生,并将在抽样中使用的一组电话号码(库)的方法;
3. 为每个电话号码制定一张在抽样中使用的呼叫清单;

4. 设计和规范问卷草稿；

5. 草拟调查员使用的引导/选择单和回撤陈述；

6. 聘用调查员和督导员，并制定访谈的具体进度和日程；

7. 进行试调查，并调整调查步骤和调查设备；

8. 打印定稿的问卷和其他表格、清单；

9. 培训调查员和督导员；

10. 进行全监控的访谈；

11. 对完成的问卷进行编辑/编码，并将数据转换成计算机可读的格式；

12. 分析数据、撰写报告等。

调查的题目一旦确定之后，电话调查方法中的下一个步骤便是进行抽样设计。抽样设计必须明确什么样的人将被抽入样本（即作为我们的调查对象）以及如何抽取这些的人（即制定出具体的抽样方案）。抽样方案的内容包括调查所代表的总体（地区或群体）、用什么样的抽样框来代表这一总体和用于生成抽样库（即那些我们用以联系被调查者的电话号码）的具体方法。在 PAPI 式电话调查中，每一个这样的电话号码都必须打印到一张张单独的呼叫单上，以使督导人员能据之对抽样过程加以控制。

问卷草稿必须先以书面形式书写出来，然后还必须对书写的格式加以规范（编排），以便于调查员用它来进行访谈。问卷题项的类型和措词在一定程度上取决于抽样方案。问卷的草稿必须和草拟的引导性的“开场白”，以及回答人选择的顺序一起进行试调查。所谓回答人选择顺序是指在抽样单位内抽取合乎调查要求的被调查人的系统方法。试调查的目的在于发现那些将要在调查中使用的各种方法和设备是否存在什么问题，以便完善和改进。此外，试调查还有另一个目的，那就是估计调查员的工作效率（如每小时完成的问卷数）和完成每份问卷平均所需的时间。试调查是任何高质量电话调查的一个重要组成部分。通常，试调查至少应实际完成 20 ~ 30 个访谈。

在试调查完成之后，我们必须撰写调查报告，这是试调查阶段一个不可或缺的组成部分。因为在最终确定我们的调查方案之前，我们对抽样方案、被调查人的选择方法、问卷和其他表格清单

所做的任何修改都必须以试调查的结果为依据。我们必须聘用督导人员和调查员，也必须确定访谈的具体日程和进度，并对督导员和访问员进行必要的培训。

在完成上述各项工作后，访谈才可以在高度控制的条件下开始进行。例如，督导员必须立即检查验收完成的问卷，监听正在进行的访谈，并据此在必要的时候给调查员提供反馈或培训。已经完成的问卷必须进行编码（即将开放式的文字答案转换成相应的数字型类别），转换成计算机可读的形式。在经过统计分析之后，我们便可以某种符合要求的形式报告调查的发现。全部过程可能只需要几天，但也可能需要一年甚至更长。

计算机辅助电话访谈（CATI）

与本书的第1版相似，本书的修订版主要面向那些使用PAPI方式的电话调查工作者。也许有人会感到惊讶，因为本书介绍的那些用于PAPI方式的方法和步骤，大多数都适用于CATI环境（注意：所谓CATI方式的电话调查，是一种访谈人员坐在计算机工作站前进行的电话调查。在这种电话调查中，访谈问卷的使用和管理，甚至整个抽样过程都在计算机工作站的控制中）。

CATI可用于控制抽样库的分布，这种控制甚至可达到为一个已经作好准备的调查员拨叫适当的电话号码的程度。CATI可以提供多种有关调查员效率的统计数据，为督导人员在人员决策方面提供帮助。CATI也可以只用于控制问卷的管理，而不控制抽样库。在向调查员提供问卷时，可以作各种类似“儿童游戏”式的复杂跳跃。它可以对问卷的题项顺序进行随机排列，也可以将前面答案编入后面的题项中。CATI还可以将文字资料同步转换成计算机可读的格式。

因为CATI具有的“高科技”形象，致使某些调查工作者，包括许多市场研究人员似乎都急于投入它的怀抱，而没有仔细考虑它的优点和缺点/费用（参见Baker & Lefes, 1988; Catlin & Ingram, 1988; House & Nicholls, 1988; Nicholls, 1988; Weeks, 1988）。CATI

不是包治百病的灵丹妙药,而是一种工具。当它恰当地应用于适当的研究时,可以改进搜集到的数据质量(有效性)。CATI 的恰当使用,不仅需要购买一些计算机及其他各种硬件和软件,而且也需要与调查设施有关的物质和社会环境有实质性的改变。

我认为电话调查向 CATI 的发展应该基于调查机构降低总调查误差这样一个愿望(Lavrakas,1991)。特别是,如果我们恰当地执行了一个 CATI,那么便应当降低:(a)与管理和处理一个抽样库有关的未覆盖误差和无回答误差;(b)与问卷题项有关的可能的测量误差;(c)与调查员语言习惯有关的可能的测量误差;(d)与数据处理有关的可能的测量误差。

在本书有关章节中,凡涉及 CATI 与 PAPI 差别的地方,特别是涉及 CATI 能使调查的有效性有所提高的那些地方,我都将予以特别的说明。对那些希望降低总调查误差的人来讲,CATI 给他们带来了很大的希望,但我们不应当把这种希望只看做是"技术上的修正",更不应当错误地认为,我们因此而不再需要那种由人掌握的,对电话调查质量进行严密控制的程序。情况恰恰与此相反:CATI 可以大大提高人在电话调查的质量控制方面的作用。

本书的内容和结构

本书的目的在于为那些认为自己并非计划和执行电话调查的专家的人,在计划和执行电话调查时提供一些帮助,并且在可能的情况下对某些从事电话调查的专家的观点作进一步的探讨。本书尤其在以下三个方面对电话调查作了比较详细的介绍:

1. 生成和处理电话调查抽样库;
2. 选择被调查者和保证得到他们的合作;
3. 组织安排调查员和督导员的工作。

本书也是塞奇(Sage)实用社会研究方法丛书中的一本。该丛书还包括福勒(Fowler,1993)的一本有关一般社会研究方法的书——《调查研究方法》(*Survey Research Methods*);福勒和麦乔(Fowler & Mangione,1990)的一本有关访谈的书——《标准化调查

访谈》(*Standardized Survey Interviewing*);亨利(Henry,1990)的一本有关抽样方法的书——《实用抽样技术》(*Practical Sampling*)。本书的主要内容是如何在电话调查中抽样、选择被调查人和督导。

本书未曾涉及的论题

以下论题本书将不作详细的讨论:

1. 调查如何应用于一个大型的研究项目;
2. 问卷题项的措辞;
3. 调查误差和调查误差估计中的统计问题;
4. 调查研究中的道德规范问题;
5. 调查数据的分析问题。

将调查方法应用于大型研究项目的知识基础

任何人,无论他采用的调查方法是面对面访谈、邮寄问卷还是电话访谈,在开始做一个调查项目之前,都必须对一般的社会科学研究方法课程或教科书(如 Babbie,1989;Cook & Campbell,1979;Hedrick,Bickman,& Rog,1993)有所了解。只有在对社会科学的含义有了足够的了解之后,才可能对调查方法的长处和局限有所理解。本书假定读者对一般的社会研究方法和社会科学已经有了一定了解,而且已经认定电话调查是一个比较好的搜集数据的方法(或者试图做出类似的决定),并希望学习更多的如何执行一个电话调查的知识。

设计改进问卷题项

限于篇幅,本书不可能就如何草拟和改进问卷题项这一问题展开任何详细的讨论。有关这一论题详细和深入的讨论可参见其他专著(如 Belson,1981;Robinson,Shaver,& Wrightsman,1991 ;Schuman & Presser,1981;Sudman & Bradbum,1982)。在使用其他教科书作为题项构建的辅助教材时,我向读者推荐通过因子分析和量表内题项一致性的信度检验,来构建多题项量表和指标(参见 Dawes,1972)。

调查误差估计

本书未对估计精确度,即调查的误差幅度的计算问题作介绍。

误差幅度这一估计值从数量上描述了调查结论的不精确程度。只要调查结论所依据的数据来自总体的某一样本或某一子集,而非整个总体,即使使用的抽样方法是真正随机的,也或多或少会存在某种不精确性。然而诚如福勒(Fowler,1993)所言,抽样误差只不过是总调查误差中的一个可能来源,我们希望读者能在这个意义上阅读和理解本书。那些水平更高的读者,不妨读一读格罗夫斯(Groves,1989)有关抽样误差方面的论著,这些论著更加明确、全面和富有挑战性。此外,我也向读者强烈推荐比墨等人(Biemer 等,1991)论述有关测量问题的著作。

调查的道德规范

福勒(Fowler,1993)和弗瑞(Frey,1989)二人都在他们的著作中谈及调查研究中的各种道德规范问题,如公布调查结果和在自愿或知情条件下的赞同的标准和规范等。本书关于这一问题的讨论,仅限于访谈中道德伦理的指导方针。当然,在许多场合,诸如这样的指导方针在一定程度上还有赖于调查研究者自身的职业判断。例如,在某人看来是一个调查员具有的“坚持不懈的劝说”的职业素养,在另一人眼中却可能被看做多少有点粗鲁或强加于人。

数据分析

介绍数据分析问题可能需要单独的教科书。调查数据本身适用于各种类型的统计分析,包括那些可将自己整合到调查中去的用于实验设计的统计分析方法。与某些人的观点相反,调查并不搜集那些本身需要用特殊的统计方法进行分析的“特殊类型的数据”。调查无非是一些方法论的实际操作技术,用这些技术搜集的数据可为各种各样的适当的统计方法所分析。

不过,在调查数据分析的最初阶段,常常会用到一种特殊的统计方法叫加权——这是一种为了更为精确地描述样本抽取的总体的事后修正。这种修正包括对选择的不等概性、未覆盖误差和无回答误差所做的弥补(参见 Groves,1989;Massey & Botman,1988)。本书不准备在此对这一论题展开论述。

借助本书制定电话调查计划

整个 PAPI 式电话调查过程由许多步骤组成。每一步骤都会

涉及一系列决策问题，本书将就以下几个方面的决策问题给读者提供比较详细的帮助：

1. 抽样库应由什么样的电话号码组成？
2. 如何处理和控制这些号码？
3. 如何选取合格的被调查人，并确保他们的合作？
4. 如何挑选和培训调查员？
5. 如何将质量控制方法应用于访谈组织和监控及调查过程的其他方面？

建立电话调查抽样库

本书第 2 章介绍如何产生一组用于抽样的电话号码。我们所需做出的第一个决定是如何生成这样的电话号码。确切地讲，我们必须确定究竟从一份清单(即一本电话号码簿)中抽取，还是使用某种方法随机地产生这样的电话号码。在做出这一决定时，我们决不能只考虑是否方便，而应该主要考虑如何才能更好地处理可能发生的未覆盖误差。

在有些情况下，如在我们的研究总体是由某些“特殊”群体(如某一社会服务机构的顾客、护士或律师)构成时，比较恰当的抽样方法是清单抽样。在这样的情况下，除非在一般人口总体中，那些我们感兴趣的特殊群体的分布密度足以达到使调查员的访谈效率保持在调查可接受的水平，否则采用随机产生电话号码进行抽样，不仅是不必要和浪费的，而且也是很不恰当的。有关这一方面的一些思考，我们将在本书的第 2 章进行讨论。

在很多场合，当研究者的兴趣在于寻找某种存在于人身上(within-person)的某种相互关系，而非对整个总体的某种特征作概括时，清单抽样也是可行的。例如，如果某一研究者希望通过电话调查了解，观众的性别和电视节目偏好之间是否相关，那么我们便可使用一本电话簿进行抽样，因为依据这样的样本得出的调查结论一般都是无偏倚的。

相反，如果我们的调查目的在于用样本统计量估计一般公众总体中的某些参数，那么以清单为抽样框进行的抽样可能就不那么合适了。一般来讲，现存的各种清单，在用作一般公众总体抽样的抽样框时，其代表性都会存在这样或那样的问题。有代表性的

清单不能说没有,但充其量也不过是凤毛麟角而已。惟一的例外可能发生在某些农村地区,在那些地方几乎所有的居民都把他们的电话号码登录在当地的电话簿上,而且外来人口进入这些地区的可能性很低,以致我们根本不必担心在最近的电话簿中会存在丢失什么新来的居民的问题。不过就大多数情况而言,对一般人口总体进行的电话调查,都应当使用某种形式的 RDD 抽样。在本书第 2 章,我们将向读者介绍几种不同的 RDD 抽样法。

怎样控制抽样库

在研究者确定采用何种方法建立电话号码抽样库(调查员将要实际呼叫的电话号码)之后,我们的注意力便应转向组织和控制呼叫这些电话号码的实际过程。人们也许会天真地以为,一个高质量的电话调查无非是让调查员一个接一个地呼叫电话号码,直至把所有列在清单上的被调查人的电话号码全部呼叫完为止。正如本书第 3 章所介绍的那样,严密控制抽样库的处理过程,是得到高质量(即有代表性)的调查结果必不可少的手段。在第 3 章我们将向读者介绍具有高度结构性(highly structured)和手选式(hand-sorted)控制 PAPI 式电话调查抽样库的方法。在介绍这些方法时,我们将对如何减少无回答误差的可能性予以特别的关注。

如何选择合格的被调查人和确保他们给以合作

我们将在第 4 章给读者介绍在访谈过程中第二阶段遇到的问题:确定适当的被调查人,并设法使他们接受我们的访谈。这时,人们也许还会再一次天真地认为,高质量的电话调查无非是让调查员将问卷中的题项一一宣读给第一个接听电话的人而已。他们不了解这种做法往往会导致性别、年龄和其他各种可能的偏倚。

在某些场合,譬如在以一份清单中列出的姓名为依据进行抽样时,抽到的那个人就是我们将要调查的那个人。这就是说被调查人的选择是由抽样过程决定的,被调查人一经抽出,便是无法更改的既成事实。但在另一些场合,电话号码只给我们提供了与被调查人生活的住户的联系途径,而在我们用电话与选出的住户进行联系之前,我们并不知道被调查人的姓名。在遇到这样的情况时,为了避免抽样偏倚,我们应当使用某种系统的方法,而不只是

依靠调查员个人的判断力来选择被调查人。系统的被调查人选择法可以提高最终样本的代表性(外部有效性,external validity),进而可以减少抽样单位内的未覆盖误差。

访谈和督导

在第5章和第6章,我们将讨论调查员的培训和督导问题。每一个优秀的电话调查员似乎本身就有一种适合做这一工作的天性,因此我们在招聘、培训和督导过程中,应当懂得如何甄别和加强他们的这一天性。尽管我们认为,在实际的调查开始之前,对调查员进行培训的确是很有必要和很重要的,但我们也认为,为了降低调查员个人的行为方式对调查可能产生的影响,从而降低测量误差,更不能忽视在调查进行过程中,对调查员进行例行的在岗培训。我们将在本书第5章对这一问题进行详细的讨论。

当电话调查比较适合我们的调查目的时,对整个电话调查过程进行控制,但将控制的重点放在电话访谈这一环节,无疑是在为保证电话调查的质量所必须采取的各种措施中最为重要的措施。访谈过程中制度化的督导程序是高质量地完成一个调查最为关键的因素。我们将在第6章对督导人员在调查过程的不同阶段的职责进行详细的讨论。

小 结

尽管市场研究人员、学术机构、媒体和政府机构一直都有做各种调查的传统,但是我们必须认识到,当今我们已经进入了一个新的时代,在这个时代,公共和私营部门的决策者都用调查搜集来的数据帮助自己进行决策。本书就是为诸如这样的"实用型"的读者而写的,其目的在于帮助那些非专业的读者学习如何操作、解释和评价高质量的电话调查。

有关调查方法的其他知识来源简介

虽然也有其他一些教科书谈到了电话调查中抽样、被调查人选择和督导等问题,但没有一本教科书,对电话调查的各种具体操

作步骤的介绍能像本书这样详细。就电话调查的写作方法而言，本书在大多数方面与其他教科书并无二致，而主要的不同之处在于本书十分明确地强调对电话调查的全过程进行全面深入监控的重要性。

此外，有些电话调查的教科书，将一些用于面对面访谈的住户概率抽样的经验方法应用于电话调查。在我看来，这种做法不仅未必恰当，有时甚至会适得其反。

正如我在前面所说，限于篇幅，本书不可能对电话调查过程中的所有方面，特别是有关问卷的设计和改进问题作深入讨论。因此我们只是在这一章，为那些希望对电话调查的基本知识有更多了解的读者，就有关电话调查的其他方面做一些简单介绍。

一般的调查方法

从理论上讲，我们应当对调查方法有一个比较广泛的了解，而不是仅仅了解某一特定的(如用电话)数据搜集方式。巴比于1989年(Babbie,1989)出版的《实用社会研究方法》(*The Practice of Social Research*)，与海德瑞克等人在1993年(Hedrick 等,1993)出版的《实用研究设计》(*Applied Research Design*)一样，将传统的调查方法与其他数据搜集方法(如内容分析、观察法等)很好地结合起来。福勒在1993年(Fowler,1993)出版的《调查研究方法》(*Survey Research Methods*)非常全面地介绍了调查研究工作者可能遇到的几乎所有问题，同时他还在该书中介绍了在其他著作中很少论及的重要论题。迪尔曼1978年(Dillman,1978)的著作《邮寄和电话调查：总体设计方法》(*Mail and Telephone Surveys: The Total Design Method*)非常详细地介绍了经作者自己整理的调查方法，其中包括邮寄调查技术。我们还向读者推荐弗瑞1989年(Frey,1989)的著作《用电话作调查研究》(*Survey Research by Telephone*)，因为该书广泛地论述了有关电话调查的各个方面。

每一个高水平的学生和调查工作者也应当了解，最近出版的一系列著作中，某些广泛论述总调查误差的各个方面的著作如《电话调查方法》(*Telephone Survey Methodology*, Groves 等,1988)、《调查误差和调查费用》(*Survey Errors and Survey Costs*, Groves,1989)和

《调查中的测量误差》(*Measurement Errors in Surveys*, Biemer 等, 1991)。

抽 样

从统计学的角度看,考克兰 1977 年(Cochran,1977)出版的《抽样技术》(*Sampling Techniques*)对调查研究中的抽样理论作了详尽的论述。而从更为实用且更容易理解的角度看,萨德曼 1976 年(Sudman,1976)的著作《实用抽样方法》(*Applied Sampling*)值得一读。同样值得我们一读的还有亨利于 1990 年(Henry,1990)撰写的《抽样方法入门》(*Practical Sampling*),它为调查业的新手学习抽样方法提供了实用的指导。最后我们要向读者推荐的是凯斯于 1965 年(Kish ,1965)出版的《抽样调查》(*Survey Sampling*),该书虽然出版年代比较早,但内容却非常全面,是一本非常有用的经典之作。我们之所以说该书出版年代较早,无非是因为该书成书于我们将 RDD 作为一种抽样方法之前。

如何设计和表达问题

虽然有效的问卷题项的构建应该属于科学方法的使用范畴,但从某种意义上讲,它又可以看做是一种手工艺术问题。虽然本书后面的某些部分将对在设计和改进新题项的过程中所必须遵循的系统步骤做一些必要的讨论,然而,问题措辞的精确程度常常取决于调查人员的专业判断力。当然,多年以来,已经有许多标准化的问题和量表面世。只要合适,我们完全可以把它们应用于我们的调查中(这就是说,文献检索常常会使我们避免做重复的工作)。但这并不是说,我们不需要对传统的量度作必要的改进。实际上,调查研究人员在打算使用以往的调查使用过的题项进行一些比较研究时,常常会因为原有题项在措辞上存在的某些不足,而处于两难的境地:如果我们对以往题项存在的不足作了"改进",那么调查得来的数据就有可能因此而丧失与以往调查数据的可比性。反之,如果不对那些不足进行改动,又会使将要进行的调查留下某些遗憾。

萨德曼和布拉德博姆(Sudman & Bradburn)曾经合作编写了三

本书,专门论述问卷设计涉及的各种问题:《调查中的回答效应问题》(*Response Effects in Surveys*,1974)、《如何改进访谈方法和问卷设计》(*Improving Interview Method and Questionnaire Design*,1979)和《提问的技巧》(*Asking Questions*,1982)。舒曼和普瑞舍二人于1981年(Schuman & Presser,1981)撰写的《态度调查中的问题和答案》读者也不妨一读,因为它对调查题项的处理问题作了详尽的介绍。最后,如果读者希望搜索有一定有效性的题项汇编,则可查阅罗宾森、赛福尔和莱特曼三人(Robinson, Shaver & Wrightsman, 1991)合编的《个性和社会心理态度的测量》(*Measures of Personality and Social Psychological Attitudes*)。

我认为那些水平较高的学生和实际工作者,还应当浏览一下由比墨等人(Biemer 等,1991)编辑成册的那本论文集中布拉德博姆和萨德曼(Bradburn & Sudman,1991)撰写的文章和其他有关问卷在测量误差中的作用的那些章节。

方法论期刊

那些对电话调查有兴趣的人还应该留意的最后一个知识信息源是在学术刊物上刊登的文献。《民意调查季刊》(*Public Opinion Quarterly*),《调查方法学》(*Survey Methodology*),《统计学》(*the Journal of Official Statistics*)和《市场研究》(*the Journal of Marketing Research*)等4本刊物,经常刊登有关电话调查和一般调查的调查方法学方面的好文章。任何一所大学的图书馆一般都会订阅这4种刊物。

练 习

1)至少另外阅读一本比较电话调查和其他调查形式之间的优缺点的书。撰写一篇两页的短文,比较本书作者和另一位作者所持的观点。

2)给当地的电话公司打电话,并尝试得到有关某一地方行政辖区内的居民电话号码的资料。然后再了解该行政区域内未在资

料清单内列出的电话号码的比例。在与公司的有关人员打交道时,务请注意有礼貌,但必须坚持不懈(如果需要的话)。无论你是否取得资料,都要写一篇短文描述一下你从公司有关人员得到的(或未曾得到的)帮助的性质(如果你得到了你所需要的信息,请给帮助过你的人写一份感谢信)。

3)编写一份执行管理一个电话调查的内容全面的大纲,其中包括一张时间进度表,大纲可从抽样方案的设计和选择开始编写。假定访谈需历时3个星期,全部访谈对象为100个。

4)从某一期刊检索并阅读一篇近来发表的有关电话调查方法的文章(即不仅仅是一篇关于调查的文章,而且是一篇关于通过电话进行调查的文章)。写一篇有关该文的长度为两页的摘要。

5)写一篇两页左右论述运用于一般公众的电话调查的总调查误差概念的短文。



建立电话调查抽样库

Generating Telephone Survey Sampling Poos

抽样库这一概念在一般的介绍调查方法的文献中并不多见。假如我们需要进行一个调查对象为1 000人的电话调查，一个没有实际工作经验的观察家也许会天真地认为，在调查中只需要与这1 000个人联系就可以了，不必再与这些人之外的其他人进行联系。但是由于各种原因，其中也包括无回答问题，实际情况几乎永远不可能像他们想像的那样。这就是说，为了达到调查要求的访谈人数，调查研究者常常需要得到比要求的人数多得多的电话号码。

这一章的目的在于让读者熟悉各种各样用于生成抽样库的方法。所谓抽样库就是一组完整的电话号码。调查员就是用这一组电话号码来得到调查最终要求完成的访谈人数（最终样本）。许多有关抽样设计的重要决策，必须在抽样库的生成之前做出。调查研究者在无法使用完备的CATI系统生成和控制这样一组用于调查的电话号码时，必须在调查开始之前，用人工方法建立一个完整的抽样库。

本章首先讨论如何在电话调查中选择有效的抽样方案。这一决定必须在确定如何生成抽样库之前做出，因为调查研究人员必须在抽样方案的引导下生成抽样库。无论采用清单抽样、随机数码拨号还是某种联合或混合方式。也不论采用人工还是计算机，抽样库的生成都必须在抽样方案确定之后。在这一章我们将讨论若干种不同的生成抽样库的方法。在本章结束时，我们将就生成的抽样库究竟应该有多大这一问题给出答案。（第2章将集中讨论如何生成抽样库，而在第3章将讨论如何处理已经生成的电话号码。）

选择有效的抽样方案

亨利（Henry，1990）为我们提供了有关抽样方案制定的全面、简明和实用的指导方针。在实际进行一项调查之前，研究者应该就有关抽样前、抽样中和抽样后三者的某些问题做出深思熟虑和明确的决定。（pp. 47～58）。这些决定应该明确的问题包括：（a）推论的总体（即调查结论意欲概括的群体、背景和时间等）；（b）目

标总体(即调查涉及的有限总体);(c)框架总体或抽样框(即操作化,常常以清单的形式出现);(d)生成抽样库的方法;(e)最终样本的大小(参见 Groves,1989,pp. 81 ~ 132;Henry,1990;Kish,1965)。

作为决策过程的一部分,调查研究者必须在使用概率还是非概率样本这二者之间做出选择。诚如亨利(Henry,1990)所言,概率样本的最大优点在于"使我们可以非常严格地考察和估计由于[它们的使用]而产生的偏倚和可能产生的误差,而使用非概率样本,[则无法做到这一点]"(p. 32)。这就是说,只有在使用概率样本时,总调查误差中那部分因抽样误差而产生的误差才可以被量化。与非概率样本不同,概率样本给抽样框中的每一单元都提供了被选择的机会(即已知的非零概率)。在事先选择的概率不相等时,调查研究员可以做出精确的调查后的修正(加权),以对不等概率进行补偿。

概率样本的类型有若干种:简单随机、系统、分层、整群和多阶等(参见 Henry,1990,pp. 95 ~ 116)。尽管我们不准备在这里明确而详细地解释这些类型之间的差别,但我们将在这一章和第 3 章,通过一些实例来讨论不同类型的概率样本的各种应用。虽然非概率样本存在很多局限性(认识这一点非常重要),但调查研究者也必须了解,并非在任何场合我们都必须要使用概率样本(认识这一点同样重要)。如果读者想进一步了解各种类型的非概率样本(方便、最熟悉/最不熟悉、典型个案、关键个案、滚雪球和定额样本)的局限性,可参阅亨利的有关著作(Henry,1990,pp. 17 ~ 25)。非概率样本一般适用于一个长期的研究项目的探索性研究阶段。

在为电话调查设计确定有效的抽样设计方案和将这一方案操作化时,调查研究人员必须对下列几个方面的问题予以考虑:

1. 调查的目的是什么?
2. 将什么人抽入样本?
3. 电话号码"丢失"的程度?
4. 可资利用的用以支持调查的资源是什么?

调查的目的

调查目的是决定抽样设计精确性的极为重要的因素。执行一项资助不当的调查常常使调查无法达到其预期的目的。在研究者和调查资助者没有(如因为不够精心或时间的限制)或不能(因为缺乏必要的知识和培训)周密地考虑他们花一笔钱究竟想得到什么时,这样的情况就有可能发生。任何一个调查决策者最低限度所必须考虑的问题是:如果一个调查并不能为我们提供有效的信息,那么它就不应当进行。

大多数人认为,电话调查(或任何其他的调查)之所以要进行,其目的不外乎以下三个:

测量总体参数

一个调查之所以进行,也许是希望确定某一变量究竟以何种水平存在于某一总体中。例如希望测定的变量水平也许是在工作中受到性骚扰的男女性的比例、安装有线电视的住户的百分比、支持还是反对增加助学税的纳税人占大多数?在这些情况下,极为重要的一点是,我们必须选择概率抽样,因为用概率样本搜集到的数据具有代表性,因而不仅是有效的,而且抽样误差也是可以计算的。我们想提醒读者的是,决定采用概率抽样本身并不能保证调查的精确性。调查研究者还必须考虑未覆盖(即使用一个遗漏了目标总体中某一部分的抽样框)和无回答(即实际调查结果中拒答率太高)问题,以及总调查误差中的其他误差源。

当调查的目的在于测量总体的某些参数时,我们应考虑的主要问题便是样本的可概括性(generalizability),或某些社会科学工作者所谓的外部有效性(external validity)(参见 Campbell & Stanley,1966;Cook & Campbell,1979)。如果调查的目的是用样本数据估计某一变量在总体中存在的水平,而选择的抽样方法却不具有较高的外部有效性,那么在大多数情况下,这种做法可能都是不恰当的。在被调查对象是以某种未加控制的和非系统的方式,就像许多非概率抽样方法所做的那样抽取时,我们就无法对抽样误差作估计。没有这样一种有关统计精确性的估计,一个调查即使不被认为是不科学的,至少也不能用它搜集到的数据对我们感兴趣

的某一变量的普遍程度进行估计。

在用计划不周的邮寄式调查和热线电话投票作调查,例如在选民的投票意向或公众对某种社会问题的态度作调查时,其结果一般都缺乏外部有效性。1984 年,在本书第 1 版写作的时候,就发生过这样一个非常引人注目的例子。芝加哥一家很有名的电台资助了一项热线电话投票调查,该调查的目的是想了解支持芝加哥市长的人多,还是支持他在市议会的反对者的人多。这次调查使用两个“900”电话号码作为热线电话,共接到电话 30 000 个左右,其中 75% 的电话声称支持市长的反对者。几星期之后,西北大学调查实验室进行了一次 RDD 电话调查(即一种概率抽样方法),并科学地确定了芝加哥居民中支持市长和他的反对者的比例分别占 40% 和 20%,而其余的人则持无所谓的态度。这些结果,与电台采用未加控制的因而也是不可靠的(即不精确的)非概率抽样方法得到的结果有非常大的差别。(顺便提一下,在下一次普选中,市长赢得了胜利,大约 60% 的选民投了他的票。)

估计多个变量之间的关系

如果一个调查的目的不是估计某一变量在总体中的存在水平,而是揭示变量之间的相互关系,那么是否采用真正的概率抽样就变得不重要了。例如当我们的兴趣在于测量惧怕犯罪与年龄、种族和性别的相关性时,能有一个异质性的样本就足够了,而不一定非要一个什么随机样本。这时,调查的目的并不是估计总体中惧怕犯罪的人的比例,而是什么人最害怕或比较不害怕犯罪。非概率抽样对这样一些类型的调查已经足够了。

评估研究

调查的第三个目的可能是评估研究(参见 Weiss,1972)。例如,为了确定一个声援曾在工作单位遭遇过性骚扰的人们地方性反骚扰的公共服务运动的效果,某一个评估者准备进行一项同组追踪研究(即一种在两个或更多的时点,对同一组调查对象进行访谈的调查方式)。在做这样的调查的第一期时,我们并不一定要从感兴趣的总体中抽取一个随机样本,因为我们要调查的是骚扰普遍程度的变化(参见 Lavrakas & Tyler,1983)。我们由此假定,感兴

趣的变量与抽入样本的人的类型无涉(即不相关)——在许多情况下,这一假定似乎是有道理的。

小 结

从以上的论述可知,调查应该采用哪种抽样方法主要取决于调查的目的。对某些调查来讲,非概率抽样已经足够精确,但对另一些调查来讲,则必须要采用概率抽样方法。确切地讲,只有在调查目的已经确定,并假定已经决定采用电话调查之后,我们才可以确定究竟使用某种形式的 RDD,还是某一现存的清单进行抽样。如果调查的目的是估计总体参数,那么 RDD 式抽样几乎总是我们首选的抽样方法。

另一方面,如果电话调查的目的是一般总体中某一较小的子集(如美国医学联合会的全体成员)的总体参数,那么采用随机/系统清单抽样则肯定更加合适。

将什么人抽入样本

究竟对什么人进行访谈这一实际问题可能也会对抽样方案的选择,尤其是会从费用的角度对抽样方案的选择有所影响。如果某一调查的调查对象必须从一般总体的某一较小的子集中抽取,那么一张清单的存在常常可以为我们提供一个精确的抽样框。例如,如果构成我们的兴趣总体的是男性护士,那么我们就需要搜集有代表性的护士清单,以建立抽样库。在区别清单中护士的性别之后,便可以以调查的名义与被调查人进行联系。相反,如果我们将符合条件的被调查人定义为一般总体中有工作的女性,那么那种随后将女性从业者从抽到的名单中分离出来的 RDD 调查,可能是我们首先应当选择的、最实用和最经济的调查方法,因为在所有抽出的住户中,可能有一半左右(至少有一个)符合条件的被调查人。

在电话抽样必须限制在某一相对较小的地理范围内,譬如某些邻里地区进行时,就可能产生比较严重的问题。在这种情况下,RDD 抽样的费用往往是难以承受的,因为调查员呼叫的住户中不符合调查规定条件的住户比例往往非常大,电话号码中的前置号

很少与邻里地区边界相一致。此外,市民也常常难以精确地确定自己究竟住在抽样的地理范围之内还是之外。在这样的情况下,如果调查的目的是估计邻里层次的参数,那么使用一本反向电话簿(即一种以住址为序排列姓名和电话号码的特别的电话簿)便可以使问题迎刃而解。遗憾的是,在我们希望搜集的数据具有很高的外部有效性时,这种方法则不可取。

电话号码丢失程度

假定我们需要对一个规模较大的公众总体进行抽样,而调查的目的又是估计与人口有关的各种参数,那么未曾登录的电话号码的比例和抽样区域内居民的流动程度,便可能对我们究竟采用RDD,还是根据电话簿或其他可得到的清单进行的概率抽样的决定会有一定影响。经验告诉我们,如果在现存的电话簿中有多至一半的居民电话未曾登录,而我们的调查目的又是有效地估计大城市中的各种总体参数,那么除了 RDD 之外,我们几乎没有其他抽样方法可以选择(参见 Survey Sampling, Inc.,1990)。在规模比较小的城市或市郊,市内电话簿中丢失的电话号码则有所减少,而在某些农村地区,电话号码丢失的可能性非常小,以致我们不必使用 RDD 进行抽样。

为了在采用电话簿还是 RDD 这二者之间做出明确的决策,我们需要了解刊登在电话簿上的住户的大致比例。有时当地电话公司能提供有关的信息,但更多的时候却不能。在后面那种情况时,调查研究人员需要对这一比例进行估计。一种估计的方法是,先确定电话簿上刊出的电话号码数,再进行多线路住户修正,最后将修正后的数字与抽样区域内的人口普查的住户统计数字作比较。尽管不存在什么公认的标准,但我认为,在丢失电话的估计数大于15%时,一般不宜使用电话簿进行抽样。此外,我还想提醒读者,除非调查的目的是估计某一变量在总体中的存在水平,否则这一问题实际上并不是什么问题。

可资利用的资源

我们最后考虑(但并非最不重要)的问题是可用于建立抽样库

和访谈的资源(包括付费的和不付费的人—时间)数量。无论是清单抽样库还是RDD抽样库,都可以用手工或计算机建立。需要建立的抽样库和最终进行访谈的样本大小都必须加以考虑。例如,要访谈的调查对象的数量是2 000,而调查的目的是研究变量之间的关系,而非估计单变量的总体参数,那么从降低建立抽样库所需人—时间的角度来讲,就用计算机还是人工方法这二者而言,使用计算机来建立抽样库也许更胜一筹。另一方面,如果最近出版的电话簿中丢失的目标总体仅为15%左右,而且调查的目的只是对总体参数值作一大致的估计,那么对那些调查经费不足的研究人员来讲,考虑到RDD需要花费较多的调查时间因而费用也比较高,用电话簿进行抽样也不失为一种较好的选择。

在考虑使用什么样的方法来建立抽样库时,我们所必须直接面对的一个问题是,如何在建立抽样库所需费用和访谈所需费用二者之间进行平衡。如果我们可以招聘到报酬较低的半工半读的学生,那么便可以用他们来搜集详细的电话信息资料,而这样的资料无疑将会有助于我们建立高效的抽样库,从而有助于降低总调查费用。因为高效的抽样库有助于减少调查员拨叫那些无用的电话号码,使调查时间大大节省,而调查时间的节省能使调查员价格不菲的报酬大大减少。相反,如果访谈费用相对比较便宜,而建立一个高效的抽样库的费用很高,那么我们便可选择一个比较容易建立的抽样库,抽样库中那些无用的号码可留待调查员在呼叫时“清除”。

对一个调查而言,有关选择什么类型的抽样的每一个决定都是很重要的。在作决定时,我们决不能武断,要考虑到方方面面的因素,并仔细地加以权衡,这就是亨利(Henry,1990)在调查计划中讨论的所谓的“创造的张力(creative tension)”。这些权衡也许要牺牲某些可能达到的目的精度。在究竟如何才能做到虽然牺牲了某些精度,但仍然能搜集到有用的(基本上有效的)数据,恐怕是见仁见智,要视具体情况而定。但是,无论做出的决定是什么,调查研究者都必须要有充分的根据。

随机数字拨号

由库珀(Cooper)于1964年首先提出的随机数字拨号(RDD)是一组概率抽样方法。这些方法为我们提供了一个进入位于抽样区域内任何一个接入电话线(即一个单独的号码,当该电话号码拨通时,电话铃声只在那一住户内响起)的住户的非零的机会。不论电话号码是否已经出版或登录,RDD都不能保证进入抽样区域内每一安装电话的住户的等概性,因为有些住户拥有多条电话线。为了纠正选择的不等概性,我们需要对多条电话线进行后抽样修正(加权),这种修正一般都在数据分析之前进行。但并非所有的不等概性都需要作后抽样的修正。修正与否,应根据调查目的而定。

虽然不是大多数,但有相当一部分美国公众的电话号码至今仍未出版(未曾刊登在一本市内电话簿上,但可以通过某种途径得到电话目录信息)或未登录(没有任何途径可得到它们的电话号码)。据抽样调查公司(Survey Sampling, Inc., 1990)最近的估计数字表明,未登录的电话号码大约为所有电话号码的三分之一。这一比例每年还会略有上升,而出于心理上的原因,我们可以预计这一数字还会继续上升,城市中尤其如此,因为大多数未登录电话号码的人一般都生活在城市。尽管我们常常假设那些未登录电话号码的住户都是高收入的白人,但实际上有少数低收入的美国人,作为一个群体,他们的电话号码最有可能未作登录。

一般来讲,抽取的样本距中心城市越远,未登录电话号码的住户的比例越低。因此近郊未登录电话号码住户的比例一般都低于市中心,远郊的这一比例则更低,而在农村地区,"未登录现象"则更少。例如,整个芝加哥市未登录住户的比例约为50%,但在它的近郊,如埃文斯顿(Evanston)这一数字就降到了25%~35%左右。芝加哥远郊地区,这一数字为20%~25%,而在郊区之外的农村地区,这一数字仅为10%。

当调查的目的是估计某一变量在一般总体中存在的水平时,

未登录的电话号码便成为代表性抽样的潜在障碍。未登录问题可能会增加调查的未覆盖误差,因为用它得到的样本并非一个包含了未登录电话号码的一般总体的随机子集。诚如前言,收入问题一般必定与未登录问题有关。而妇女、单身成年人和那些教育程度较低的人更有可能成为未登录者。毋庸置疑,对犯罪问题的担心程度与未登录电话号码的多少直接相关(Lavrakas 等,1980)。此外,我们还发现未登录住户的无回答率也可能会略微高一些(Drew,Choudhry,& Hunter,1988)。

RDD 降低了由未登录电话号码引起的潜在的未覆盖误差。在调查研究人员掌握了抽样区域内电话线路的前置号码(即市内电话号码的前 3 位)之后,便可以用各种方法在前置号码后加上后置号码得到 7 位的市内电话号码(跨地区码的电话调查必须使用 10 位的电话号码)。当然这些号码也可能会有一部分空号。如果这些号码不是空号,则不论这一号码是否已经登录,它都有可能接通一个住户或非住户。

如果我们已经决定采用某种形式的 RDD 作抽样,那么就可以用手工或用计算机建立将要由调查员处理的号码库。当然,用计算机来做这一工作是上乘之选,但并不是一定要这样。从理论上讲,任何用计算机建立 RDD 抽样库的方法,都可以用手工方法完成。然而实际上,由于太浪费时间而不用手工方法。

建立大多数 RDD 抽样库需要的各种准备工作

RDD 一个非常实用的优点是,它不需要有一个抽样赖以进行的完备而详尽的抽样框。这就是说它不需要一张罗列了抽样区内所有可能的电话号码的清单。如果我们想要一个含抽样区内所有可能的电话号码的抽样框,RDD 可以轻而易举地为我们复制一份。如果我们打算建立一个非常完备的抽样框,那么 RDD 产生的抽样库就是这样一种抽样框的十分完美的复制品。因为在这一抽样库中,抽样区域内的所有可能接通的电话号码抽入样本的可能都已经产生。

搜集前置号

大多数用 RDD 法建立抽样库的第一步是,搜集制定可在登录抽样区域地理范围内接通的电话的前置号码的完整无遗的清单。而这一任务通常不是那样简单明了和很容易完成的。有时抽样区域的边界和在抽样区内接通的电话的前置号码完全一致,有时二者则略有差异。但有些时候,二者之间存在比较大的差异。在遇到后面这样的情况时,我们也许不应该采用 RDD 抽样,因为费用过于昂贵。而昂贵的抽样费用有可能抵消调查本身的价值。在前置号的边界相当接近抽样区域的边界时,用一些对地理位置进行筛选的问题(geographic screening questions)将那些居住在抽样区域以外地区的住户排除在外,或采用相反的方法,即将原定的抽样区域缩小或扩大,使二者归于一致。正如我们一贯强调的那样,我们不应武断地做出有关决定。调查研究者的责任是确定调查目的可以在多大程度上容忍诸如这样的变换。

我们建议那些准备在全国、全州或地区(在整个一个区号码,或若干个区号码之内)范围内进行的电话调查,使用美国电话电报公司(AT&T)的所有这些地区码和前置号码数据库的最近的软盘(或磁带)。因为在美国和它的周边国家,这些数据库总是在不断地变动。在 1992 年,这一数据库可以在位于新泽西州莫里斯顿市的贝尔中心(Bellcore, Morristown, New Jersey)购得,价格为 400 美元左右。数据库共有地区码和前置号的组合 50 000 个左右,尽管所有这些号码组合并非都用于住宅的电话线路。

那些不准备进行覆盖整个区号(如调查 10 个相邻的市郊或 3 个县等)的 RDD 调查,有关前置号码的信息资料则必须靠人工搜集。而我们期望可从市内电话公司得到的有关前置号码信息资料的帮助则因地而异。比较简单易行的方法是认识某个在电话公司工作的人,要求他给我们提供我们要寻找的有关信息资料。除此之外,研究人员个人在寻求电话公司给予帮助的过程中,若能做到锲而不舍则必将有所收获,若遇难而退则可能一无所获。

如果电话公司最终仍然不愿提供帮助,那也没有什么关系。因为即使使用一本反向电话簿,甚至使用一本普通的电话簿我们也可以搜集到相当可靠的,本应从电话公司觅得的信息资料的估

计值。比较而言,使用反向电话簿更为容易些,因为编辑出版这样一种电话簿的目的就是要按地理位置的顺序来排列电话号码。此外,有些反向电话簿也提供一些有用加总统计数字,如登录的分不同前置号的住宅用户数。

在没有反向电话簿可资利用时,完成这一任务便会更加费时,有时甚至根本无法可施。在搜集到适用于抽样区域的所有电话簿之后,调查研究人员便可以在理论上确定前置号适用的区域是否与抽样区域的边界匹配。我们根据从电话簿上找到的地址,将对应的前置号系统地在地图上一一标明就可以了。根据抽样区域面积大小的不同,完成这一任务的时间也不同,有时甚至会相当漫长。实际上,除了某些面积相对较小的地方区域,这种方法一般是很不实用的。

确定每一前置号码的线路数

正如下面所要详细介绍的那样,掌握了解抽样区域内每一前置号码的正在实际运行的入户线路数是很有用的。如果每一前置号码在这一方面的信息资料的精确性相同,那么我们就可以用它按每一前置号码在抽样区域内的比例产生随机电话号码。采用这种方法最终建立的抽样库是按前置号码分层的。最理想的情况是,我们可以在抽样区内的电话公司中找到这些资料。即使这些资料必须依靠电话簿才能得到,我们仍然可以从电话簿页码的代表性样本中系统地抽取前置号码。然后再以此为根据估计将要使用的不同前置号码内住户电话线路数的相对比例(后一种方法假定跨前置号码未登录的电话号码的比例是恒定不变的)。如果这一假定不成立,我们就不能使用这一方法。

例如伊利诺斯州埃文斯顿市有 8 个住宅前置号码。而从埃文斯顿市的电话簿中抽取的电话号码样本可知,2/3 的入户线路所使用的前置号码是其中的 3 个。在采用 RDD 法建立埃文斯顿市的抽样库时,(从调查员处理的时间角度看)那种反映前置号码分布的方式是最有效率的。这种方法在用于相对较小的地理区域的抽样时相当方便。

确定未运行后置号码的范围

有关未运行后置号码范围(未运行或含非入户线路的电话号

码分布的数字范围)的信息资料也可以用于提高改进 RDD 抽样库的效率。如果不是马上可以从电话公司得到这类资料,那么我们就可以使用反向电话簿。只要我们仔细地扫描每一前置号码登列的数字顺序,便不难发现每一前置号码未运行的后置号码的数字范围。有效地使用这一方法的前提是,电话公司分配给未登录用户的号码不是来自某几组特定的后置号码。换言之,电话公司的做法应当是,先将一个可以使用的电话号码分配给一位新顾客,然后再问他/她是否愿意出版或登录这一号码。因此只要用肉眼扫描一下反向电话簿上登录的电话号码的数字顺序,我们便能系统地确定那些正在运作的后置号码的数字范围。我向读者推荐寻找这类数字范围时使用单位 100——这就是说,调查研究者应当确定,每一前置号码的后置数码在范围 0000 ~ 0099,0100 ~ 0199,0200 ~ 0299……是否存在。

当抽样区域内含有的前置号码数以百计时,这一工作便不再简单了。诸如这样的信息资料将明显减少 RDD 抽样中的未运行号,而未运行号的减少无疑会减少访谈的时间和费用。如果得不到反向电话簿,那么使用一本普通的电话簿也可以,但这时完成这一任务所需要的时间则难以计数。究竟是否要搜集诸如这样有关前置号码的资料信息,仍然取决于调查研究者自己。例如,如果有低费用的或“免费”(即无直接费用,只需一定的时间和补助)的学生可用于搜集这样的信息资料,那么通过搜集这样的信息资料来提高抽样库的效率无疑值得一试。

如果我们想在同一地理区域内进行若干次调查(如每年作一次),那么便值得在初次调查时花一笔钱搜集尽可能多的有关该地区的前置号码的资料信息。重要的问题在于,我们必须根据抽样区域内居民的流动情况,定期地(如一年)更新有关分配给每一前置号码的住户比例和未运行后置号范围的信息。

简单 RDD 抽样和多阶 RDD 抽样

正如下面我们将要更为详细地讨论的那样,有时研究者也许会遇到只能得到有关抽样区的前置号码,而无法得到其他任何信息资料的情况。在这样的情况下,我们可以使用以下两种方法中的一种来进行抽样。第一种方法,我们称之为简单 RDD 抽样。其

做法是我们在每一前置号后面，加上数目相等的随机后置号，然后再在访谈时将那些无效的号码（即那些在一个一般公众调查中不入户的电话号码）从抽样库中清除。而在使用第二种方法时，使用最多的是二阶抽样法，即抽样库的建立可分为两个阶段。在第一阶段，先用简单 RDD 方法产生数目相对较小的电话号码。然后我们呼叫所有这些电话号码，而其中那些叫通的住户号码将用于在第二阶段建立完整的抽样库。

有关二阶抽样究竟应该怎样操作的问题我们将在后面进行讨论，但是在这里这一问题已经出现，因为研究者现在必须就抽样库的建立，究竟需要搜集多少信息资料这一问题做出决定，而这一决定在很大程度上取决于搜集这类信息资料和访谈所需的实际费用。如果访谈所需的费用高于搜集抽样库资料的费用，研究者可能应当选择建立最为有效的抽样库。如果情况相反，那么研究者则应选择建立虽然比较容易建立，但未必是最为有效的抽样库。让那些“比较廉价的”调查员（如无报酬的学生），通过拨打比例相当高的无效的电话号码来清洁抽样库。（需进一步考虑的问题是，相对于简单 RDD 法抽样而言，二阶抽样的抽样误差也比较高。）

人工建立 RDD 抽样库

虽然人工建立 RDD 抽样库的方法很少有人使用，但如果我们的讨论从这一方法开始，将有助于我们对抽样库建立问题的理解。而我们之所以以这样的顺序来展开我们的讨论，是因为相对于更快和更有效的用计算机建立而言，手工建立抽样库的方法在方法论意义上更清晰和更为简单，因此也更容易为初学者所理解和掌握。

有两种可以用于人工建立 RDD 抽样库的方法。第一种方法是使用随机数码表（我们可以在大多数统计教科书的后面找到这些表格）。我们可以先从表中选择一连串数字作为后置号码，然后再在它们前面加上用于抽样区的三位前置号码。第二种方法是数字加入（added-digit(s)）法，即在直接从电话簿上抽到的电话号码上加上一个或几个数字。严格地讲，这种方法可能不能满足为抽样区域内每一可能的电话号码都提供一个非零选择的可能性的

要求。

随机数码表法

我们先用一个例子来介绍最为简单的建立 RDD 抽样库的方法。在使用这一方法时，研究者所掌握的惟一的信息资料是抽样区域内的前置号码。在将搜集到的前置号码登录在清单上之后，我们只需一次一个地将随机数码表上相邻的 4 位数字加到每一前置号码之后即可。假如我们在随机数码表上找到了一串数字：547196353826，并且在本例的抽样区中只有 3 个前置号码：864、866 和 869，那么我们首先找到的 3 个电话号码便是 864-5471、866-9635 和 869-3826。依次类推，将这一过程进行下去，直至产生的电话号码足够为止。（估计抽样库是否已经足够大的方法将在本章稍后部分讨论。）

如果除了抽样区的前置号码之外，研究者还掌握了有关抽样区的前置号码的电话接入线路数的相对比例的完整而精确的资料信息，那么这一信息资料便可降低随后将要进行的访谈时间。在上述 3 个前置号码中，假如前置号码 864 含有抽样区中 20% 的进入住户的线路，而 866 和 869 各含有 40%，那么在用随机数码表建立抽样库时应以 864，866，866，869，869，864，866，866，869，869……这样的顺序，将一串串 4 位数字加到每一前置数码之后，以使生成的电话号码数能反映出前置号在抽样区的比例（即 2∶4∶4）。

如果除了上面两种信息之外，我们还掌握有关未运行后置号码范围的精确的和完整的信息资料，那么这一信息同样也可以用于提高 RDD 抽样库接入运行号码的效率。我们再一次使用前面的例子，即使用与前例相同的 3 个前置号码，假定前置号 864 中运行的后置号码的范围为 2000 到 3999；866 中为 5000 到 7999；而 869 中运行后置号码的范围为 0000 到 0999 和 4000 到 4999，那么在建立抽样库时，我们只需从随机数码表中，选取一串串长度为 3 位的随机数。然后将这些 3 位随机数加入如下的前置号码和已知在运行的前导后置数：864-2，864-3，866-5，866-6，866-7，869-0 和 869-4。萨德曼（Sudman，1973）把这种处理方法称之为逆抽样法（inverse sampling method）。

读者必须注意的是，不论我们使用何种方法建立 RDD 抽样

库,如果需要使用有关前置号的各种信息,那么不仅要求在抽样库中的所有前置号码的资料信息必须都存在,而且各个前置号的信息的精确性也都必须相同,只有在这两个条件都能得到满足时,我们才可以在抽样库的建立过程中,合理地使用有关前置号的信息。其原因在于如果没有有关所有前置号码精确性相同的信息资料,研究者就无法精确地计算每一前置号码所需产生的号码数的相对比例。不仅如此,要能恰当地使用这一额外的信息资料,我们还必须假定,每一前置号码内未登录的号码数是大致相等的。一般来讲,对于许多抽样区域而言,这一假定可以认为是合理的。而只要上述假定中的其中一个不能得到满足时,我们认为还是选择简单RDD 抽样(或下面将要介绍的二阶抽样)为好,而把有关问题留给调查员在处理抽样库的电话号码时处理,这样可使前置号的抽取比例比较适当。

因为 RDD 调查具有自加权的特点,所以它所采用的抽样方法是正确的。这就是说,在理论上讲,调查完成的样本中的这些前置号码的比例,将与它们在抽样区内的比例一致。在我们掌握抽样区内的前置号的可比信息资料时,这种所谓的自加权处理的效率便会因此而得到加强。如果诸如此类可比信息资料不存在,那么还是不要“画蛇添足”为好。这一原则对有关空置前置号的信息资料的处理也同样适用。除非我们能搜集到抽样区内所有前置号的完整的信息资料,且所有的前置号中的未登录的电话号码的比例大致相等,否则还是采用那种简单明了的抽样方法为好,即按相同的比例给每一前置号分配等量的电话号码,并在每一前置号后加上个数相等的随机号码。

此外,我们认为对所有的 RDD 抽样法还必须要加以说明的一点是:我们有可能把由因掌握每一前置号码中运行号码的比例,和每一前置号中空置后置号码的范围的信息而产生的效应合并在一起。用手工来做这一工作虽然是非常复杂的,但是是可能的。而在前置号码数不是太多的时候,这种做法是可行的,当然它需要一定的时间和精力。

数字加入法

兰德勒和班克斯(Landon & Banks,1977)提出了一种检验建立

电话调查抽样库的效率的经验方法。在使用这种常规的检验法时,我们必须先从一本电话簿中抽取"种子"号码,然后再将某一固定或随机的数字(可以是2位,也可以是3位)加到每一种子号码的后置号码上,生成新的电话号码,最后再在抽样库中使用这些新的号码。用这种方法产生的新号码可能已经登录,也可能还没有登录。列普科沃斯基(Lepkowski ,1988)认为:"要使未登录的电话号码有被抽入样本的可能性非零,我们必须假定未登录号码均匀地分布于登录的号码之中。"(p.93)。他认为这是"一个不可能成立的假定"。虽然我并不完全赞同列普科沃斯基的观点,但我支持他对这一问题所持的谨慎态度:尽管就应用而言,这些方法都比较简单,但它们无法确定生成的单个电话号码被选择的概率。尽管有了这样的考虑,但研究者所面临的处境仍将迫使他在可行性/及时性和总调查误差略有增加的可能性这二者之间做出选择。正如本书始终都在提醒读者注意的那样,这些决策无非是调查研究者经常遇到的许多实际决策(权衡)中的几个而已。

用于加数的方法可以是多种多样的。例如"加1法"便是可以使用的方法之一。在使用这一方法时,假如我们从一本含抽样区域的电话簿上随机或系统地抽取一个电话号码869-5025,在这一号码的后置号码的最后一位加上一个1,便得到了数869-5026,这一数字便是随后进入抽样库的数字。另一种可以使用的方法是加11法,或加一位或二位的随机数法。即在每个从电话簿上抽出的种子数上加上11(因此而得到869-5036),或一位或二位的随机数。

当我们必须在很短的时间内用手工建立一个比较小的(例如含1 000个以上电话号码的)抽样库时,这些抽样方法特别有用。与简单RDD抽样相比,这些方法建立的抽样库增加了将要拨叫的正在运转的电话号码的比例,从而缩短了访谈周期,加快了整个访谈的进程。兰德勒和班克斯(Landon & Banks,1977)在两个研究报告中指出,加数法可把抽样效率提高30%左右。这一效应与前述事先掌握有关未运行的后置号码的范围,以使抽样产生的电话号码尽可能在运转的后置号码范围之内的方法不相上下。

在使用加数法时,必须首先搜集一套全面而完整的电话簿,然后确定从每一本电话簿中应该抽取的种子号码数。当然最为简单

的情况是只有一本而不是一套电话簿。在估计了达到最终希望完成的访谈的样本量所需的抽样库的大小之后，调查研究者应该将这一估计值与刊登在一本或几本电话簿上的电话号码总数进行比较，然后用随机或系统方法，在整个电话簿的页面上选取种子号码。

兰德勒和班克斯(Landon & Banks,1977)担心，如果电话公司将未登录的电话号码集中在某一范围的后置号码内，那么加数法抽样可能会产生偏倚(参见 Lepkowski,1988)。如果这种情况真的发生了，那么以刊登在电话簿中的种子号码为依据的电话号码样本就有可能产生相当大的未覆盖误差。换一句话说，就未登录的电话号码而言，拨叫的电话号码库是有偏倚的。尽管有这样的担心，但我并不认为有任何一家电话公司实际上就是以这样的方法来分配未登录电话号码的。相反，电话公司的做法是先分配号码，然后再让顾客决定自己的电话号码的处理方式，在刊登、不刊登和不登录这三者中进行选择。尽管如此，为了保险起见，任何人在准备采用加数法之前，为谨慎起见，最好还是事先了解一下电话公司是如何处理未登录号码的。

如果一个调查研究人员准备进行一个一次性的(如为硕士论文搜集数据)电话调查，并且无法确定这种搜集数据的方法在不久的将来是否还会再次使用，那么如果样本库不是非常大(有几千个电话号码)的话，那么人工建立 RDD 号码库的方法是比较合适的。不过，如果必须处理的号码库非常大，或调查研究人员希望定期在一个抽样区内多次进行电话调查以搜集数据，我们则强烈地推荐使用计算机来建立 RDD 抽样库。

用计算机建立 RDD 抽样库

所有用计算机建立 RDD 抽样库的方法，都是从搜集编制在抽样区内运行的全部前置号码的穷尽无遗的清单开始的。正如我们前面所讨论的那样，搜集有关每一前置号码的其他额外的信息资料，特别是搜集有关(a)每一前置号码的用户数；(b)每一前置号的空号范围的信息资料的成本效益是相当高的。但我们想提醒读者注意的一个重要问题是：除非那些额外信息的精确性和可用性对

所有用于抽样的前置号码大致相当,我们才可以在抽样中使用它们,否则就不可以使用。

我们可能得到的前置号的信息资料有三种类型,而研究人员可能遇到的情形则有四种:

1. 只有前置号码清单;
2. 有前置号码清单和每一前置号码的用户数;
3. 有前置号码清单和每一前置号码的空号范围;
4. 有所有三种类型的信息资料。

只有前置号码

最简单的情况是只有抽样区内的前置号码的资料信息。在遇到这样的情况时,我们只需将计算机程序产生的随机后置号码等量地分配给每一前置号。当然我们必须首先确定要求计算机产生的随机号码的个数,然后用编好的程序,令计算机一个接一个地生成随机数,并将它们等比例地分配给各个前置号,直至产生的随机数数目达到事先规定的个数,即达到抽样库要求的电话号码数。例如,如果我们估计抽样库需要的 RDD 随机号码数为 1 000 个,而样本区内共有前置号码数 10 个,那么将 1 000 作简单的等分,则每一前置号码都将配给 100 个惟一的四位随机数串。依据使用的打印机打印速度不同,如下面所列的 BASIC 程序,将在一部 DOS 系统的计算机上,在几分钟内产生并打印出 1 000 个号码(在 DATA 语句中列出的 10 个前置号码,每个各 100 个)。这一程序运行产生的结果也可以存放在一个 ASCII 文件中,等以后需要的时间再打印。

```
 5  RANDOMIZE TIMER
10  FOR I = 1 TO 10
20      READ PREFIX(I)
30      FOR X = 1 TO 100
40        SUFF1 = INT(RND * 9)
50        SUFF2 = INT(RND * 9)
60        SUFF3 = INT(RND * 9)
70        SUFF4 = INT(RND * 9)
80        LPRINT PREFIX(I);"-";SUFF1;SUFF2;SUFF3;SUFF4
```

```
90      NEXT X
100   NEXT I
110   DATA 251,256,328,475,491,492,570,864,866,869
```

因为用计算机(甚至用随机数码表)生成随机数字串,都会存在产生两个完全相同且相配于同一前置号码的数字串的小概率。换一句话说,如果计算机产生的后置数字串为4567,且把它相配于两个不同的前置号码(如251-4567和869-4567),这就不会产生什么问题。但是,有时两个恰好相同的电话号码偶尔也会出现。在用手工建立抽样库的情况下,用人工来检查是否有重复的号码几乎是不可能的,但用计算机来进行这一工作则是很容易的。因此,只要可能,我们就应该在程序中加入一个子程序,用它检查重复的电话号码(例如,在使用很广泛的统计软件SPSS中,我们可用"lag"命令来检查重复的号码。参见LeBailly & Lavrakas,1981)。

除清除抽样库中的重复号码之外,我们还可以编写令计算机进行其他的清洁工作,如令计算机清除所有以00,000和/或以0000结尾的号码的程序。这些号码一般都不是住宅号码,因此把它们清除,既可以缩短访谈过程,又不会降低抽样库的效度(即不会增加未覆盖误差)。但在做这样的清除的时候不仅要非常小心,而且只有在核对了反向电话簿,并对假定的有效性予以确认之后方可进行。因为企业,由于其本身的性质所决定,一般都会把它们的电话号码刊登在电话簿上,因此我们只要浏览一下反向电话簿,就可以弄清所有以多个零结尾的电话是否几乎都是企业的电话号码。如果不是,那么诸如这样的例行清除便是不可行的,否则势必会导致未覆盖误差的增加。

未运行后置号码范围

如果未运行的后置号码的范围已在我们的掌握之中,那么这些信息资料也可编入计算机程序。完成这一工作的方法有两种。第一种已在前面提到,如果我们可以确定未运行的后置号码的数字范围为1 000(即1000~1999或3000~3999),那么就可以把资料信息编入计算机程序。第二种方法是得到四位数字号码的信息资料(即三位前置号码和一位后置号码的前导号码)把它们用DATA命令读入,并用程序令计算机产生三位的随机数字串。我们仍

使用前面的例子，假如抽样区有三个前置号码，而每一前置号码的运转后置号码的值域比较窄：从 864-2000 到 864-3999；从 866-5000 到 866-7999；从 869-0000 到 869-0999 和从 869-4000 到 869-4999。在这样的情况下，下列数字串：8642，8643，8665，8666，8667，8690 和 8694 将被作为 DATA 读入。而我们可以用程序令计算机在每个读入的四位号码后，加上三位的随机数字串。用这种方法建立的 RDD 抽样库，它的电话号码中的后置号码，必定都集中（分层）在那些已知的包含正在运转的住户电话的后置号码范围之内。

入户线路数

最后，如果调查研究人员已经知道每一前置号码的电话入户线路数，那么 RDD 抽样库中的号码便可以以前置号码分层。为了行文简便起见，我们还是假定后置号码的空号未知。如果已经确定前置号码 864 有运行的电话号码 4 000 个，而 866 和 869 分别有 3 000 个和 1 000 个，那么我们便可以先计算抽样区中每一前置号码的相对频数（比例）。在总数为 8 000 的入户线路中，50% 经过 864 入户，而通过 866 和 869 入户的比例分别为 37.5% 和 12.5%。我们可由此确定，如果研究者估计调查员可能需要处理的 RDD 电话号码为 1000 个，那么抽样库中的电话号码带前置号码 864，866 和 869 的数目应分别为 500 个，375 个和 125 个。

为了使在 DOS 操作系统下的计算机中能完成这一任务，我们只需将前面列出的程序修改成如下的程序便可：

```
5   RANDOMIZE TIMER
10  FOR I = 1 TO 3
20    READ PREFIX(I)
25    READ NUMB(I)
30    FOR X = 1 TO NUMB(I)
40      SUFF1 = INT(RND * 9)
50      SUFF2 = INT(RND * 9)
60      SUFF3 = INT(RND * 9)
70      SUFF4 = INT(RND * 9)
80      LPRINT PREFIX(I);"-";SUFF1;SUFF2;SUFF3;SUFF4
90    NEXT X
```

```
100  NEXT I
110  DATA 864,500,866,375,869,125
```

如果我们不仅掌握了每一前置号的入户电话线路数，而且也掌握了有关每一前置号码的后置号码的空号范围，那么这一信息资料同样也可被编入计算机程序。当然要实现这一目的，我们也必须像前面所做的那样，对原有程序做一些修改，增加一个子程序，以反映每一前置号码在最终的抽样库中所占的比例。

多阶 RDD 抽样库

米托夫斯基（Mitofsky，1970）和瓦克斯伯格（Waksberg，1978）在探讨比较复杂的 RDD 抽样方法方面做了许多开拓性的工作。他们创造的方法不仅对建立高效的大容量的抽样库具有很高的实用价值，而且可能在抽样误差略有增加的前提下，使访谈费用显著降低。正因为如此，他们创造的方法为许多大型调查公司所广泛使用。这种方法一般可分为两个阶段：在第一阶段，用简单 RDD，在抽样区内选定一批数目相对较小的住户电话号码（一般为 50 ~ 100 个）。然后把这些号码作为第二阶段的种子（算子）。在进行一个大规模的调查（如全国、地区和州规模的调查）且不存在与调查区域匹配的完整的前置号码的清单，而要重新搜集这样的信息资料又令人望而生畏时，这样的抽样方法就更加有吸引力。尽管采用这种方法可能会使调查的抽样误差略有增加，但它的使用不仅十分方便，而且还有较高的效率（参见凯思（Kish，1965）、格罗夫斯（Groves，1989）、列普科沃斯基（Lepkowski，1988）和亨利（Henry，1990）等人有关设计效应的讨论）。在对这样一种交换的得失进行评价时，格罗夫斯说："尽管调查统计值的精度因聚类样本而有所丢失，但所有美国联邦政府进行的大型调查和几乎所有大型学术性的及商业性的住户调查，都无一例外地采用聚类设计。"（p. 260）

一个全国性调查的实例可能会有助于我们理解如何用二阶 RDD 法建立抽样库。在做这样的调查时，调查研究人员首先应该得到一份可靠的包括全国所有住户电话的前置号码的清单。得到这样的信息资料最为容易的方式是，购买贝尔中心的计算机化了的全美地区码和前置号码的数据库（1992 年，价格在 400 美元左

右)。在20世纪90年代早期,所有可能生成的用于美国的RDD号码,约有1/4为住宅号码。我们根据某些抽样专家的意见,事先确定要使用的串(cluster)的数目(在第一阶段确定的一个个单独的住户号码数),一般它应该在50到100个之间。在第一阶段,我们可能需要建立一个容量4倍于将要使用的串数的初级抽样库。一个使用贝尔中心数据作为抽样清单且准备用100个在第一阶段确定的串,作一个最终样本量为1 100的调查研究人员,可能需要从贝尔数据库中约50 000个的地区码和前置号码的组合中,随机地选出400(100×4)个。然后再将随机的后置号码加到每一选出的组合号码上,得到400个10位的电话号码。

然后,调查员应该拨打这400个号码,以确定哪些号码可进入住户。一般我们可望得到100个左右的入户号码。实际操作过程是很简单的,但研究者在令调查员开始拨打这些电话之前,应该先熟悉一下有关材料(实例请参见亚历山大(Alexander,1988)、博克海默和列瓦索汉(Burkheimer & Levinsohn,1988)、梅森和埃莫曼(Mason & Immerman,1988)等人的著作)。有时,调查员只要用一个简短的“筛子”式问题,便可确定接电话的用户属于什么性质,但有时调查员则可能要对打通的每一个电话号码进行详细的询问,才能最终确定接电话的用户的实际性质。不论第一阶段使用的所有电话号码是否可被归入住户类(包括拒访和住宅应答机),这一工作都可以通过访谈来完成,但对那些从来没有人接答的电话号码,研究人员则必须花费一定的时间和经费,与市内电话公司联系,以确定这些号码是否是住户电话号码。(例如NUSL于1992年进行的二阶全国性RDD调查,在第一阶段大约有3%的电话号码,尽管调查员试呼次数在20次以上,但仍然无法确定它们的性质,必须与市内电话公司联系,方能确定。在这一调查中,为了确定抽样库中所有号码的性质,第一阶段进行了两个星期。)

在全国性调查的第一阶段中确定的住户电话号码,是全国所有住户电话号码中有代表性的简单随机样本。而在目前介绍的例子中,因为最终计划完成的样本容量为1 100。假如我们共有11个调查员参加第二阶段的工作,由他们来处理完成所有100个第一阶段得到的串,最终产生调查要求的1 100个样本。那么我们就应

该这样做:取出每个在第一阶段得到的住户号码,然后再随机地取代其后置号码中的最后两位(至少做 11 次)。如此这般,一个串一个串地反复进行下去,直至我们希望完成的所有号码全部得到为止。因为每个串至少在第二阶段需要生成 11 个电话号码,所以最初建立的第二阶段抽样库一共可以生成 1 100 个单独的电话号码。以往的经验告诉我们,可以与住户接通的电话号码,约占全部第二阶段电话号码的 60%(Lepkowski,1988)。将这一数字与第一阶段 25% 的住户"击中率"相比较,不难看出这一方法可以明显减少访谈时间,也就是说,第二阶段的"击中率"是第一阶段的 2 倍多。在第二阶段,每个无法接通的电话号码,都将被该串中新产生的号码所取代。

在实际调查中,第二阶段的抽样库的建立是一项很复杂的工作,因而最好设法使用计算机数据库,以便经常更新。此外,每一串最终接通的电话号码数目可能不完全相等(这一问题可以通过后抽样的加权加以修正)但并非所有的串都可以产生我们所希望的最终可以接通的号码(参见 Lepkowski,1988)。我们希望那些打算使用多阶 RDD 抽样方法的读者,通过进一步阅读其他的有关著作(如 Groves 等,1988;Potthoff,1987),并在阅读过程中结合自己特殊的应用目的,与他人在运用这些方法时积累的丰富经验展开交流讨论,深入地了解这一方法在实际使用和理论阐述之间存在的细微差别。

以电话簿和清单为依据的抽样库

由于调查目的的不同,在很多时候,RDD 并不一定能满足所有调查的需要。读者必须记住,虽然在各种可以用于从一般总体中选取有代表性的(有电话)住户样本的方法中,RDD 是一种避免未覆盖误差的比较好的方法,但是如果我们可以得到兴趣总体的精确清单,那么抽样工作也许还是依据清单为抽样框进行为好。

问题在于,在很多时候调查研究者往往无法得到他们感兴趣的清单。弗瑞(Frey,1989)曾对与电话调查中清单抽样有关的问题

进行过讨论。这些问题包括：禁止进入线路、不合格的清单、重复登录的清单，以及不正确和有遗漏的清单等。例如，假如每年美国约有 1/5 的住户搬家，那么许多清单登录的可能都是那些过时的住户电话号码。

一个有关用清单进行抽样的例子可以是对某一专业团体（如退役军官组织）成员的电话调查。在这样的调查中，该团体拥有每一成员的家庭和/或单位电话号码的清单，而这样的清单便可作为抽样框使用。另外一个有关清单抽样的例子可以是对目前在大学注册的大学生的电话调查。在这一例子中，来自注册办公室的清单（抽样框）可能定义了整个兴趣总体。正如我们前面所提到的那样，在美国某些农村地区很少有人不登录自己的电话号码，且人口中新进入人口的比例也相当低。在诸如这样的地区，我们可以市内电话簿为根据来选取抽样库电话号码。但在整个抽样过程中，研究者始终不要忘记对未覆盖误差的监控。

用清单抽样建立抽样库的过程并不复杂，我们可以随机地或系统地从清单中选取抽样库电话号码。如果可以用系统方法从整个清单中进行抽样，一般则不必使用随机方法。尽管如此，也许仍然会有人希望采用随机方法，因此有必要对这些方法都作一些适当的介绍。

随机方法

假如我们要在一个拥有 1 800 名教师的校区，对 200 名教师进行有关工作条件满意度的电话调查。那么最简单的做法是用程序在校区计算机上，随机地抽取 200 名（或更多）教师作为抽样库。（请读者务必记住，几乎在所有的电话调查中，抽样库都必须大于计划得到的最终样本量，这在一定程度上是因为拒访的缘故。但在这一例子中，我们假定所有抽入抽样库的教师的电话，不仅都可以被接通，而且选中的被调查人都配合参与调查。）如果无法使用计算机，我们可以用随机数码表，并可以采用多种方式来完成这一工作。譬如一种方法是，用两位的随机数字串；在一张按姓氏字母次序排列的全部 1 800 名教师的电话号码的清单上“向前计数”，一个接一个地选出被调查的教师。例如随机数码表上列出的数字串

为0417565439,那么在按姓氏字母排列的清单上的第四个人,便是第一个进入抽样库的人。然后再向前数17个人,便是第二个进入抽样库的人,第三个人则应当再接着向前数56,如此这般,反复进行。用这样的方法从1 800个人的清单中抽取200人建立抽样库,可能需要在清单上从头到尾重复五六次。但究竟需重复多少次,则需视具体情况而定。如果从头到尾抽取一次,可抽出的人数的均值为36,那么完成200个人的抽样需要从头到尾进行5.5次。

系统方法

一般来讲,从减少抽样误差的角度看,更容易且更受欢迎的方法是从一份清单上作系统抽样(参见 Henry,1990)。我们仍以从一个有1 800名教员的校区抽取200名教师作为调查对象的电话调查为例。在这一例子中,全部1 800名教员中,将有1/9被抽入样本。首先我们从1到9这9个数字中,随机地选取其中一个作为起始数,并从以姓氏字母次序排列的清单中的第一个名字开始往下数,直至数到与选出的起始数对应的那一个位置为止,在那一位置上的那名教师就是第一个被选入样本的教师。在第一个教师被抽出之后,我们再从这一位置开始,接着往下选,每隔9个选一个。如果随机起始数为2,那么清单上的第二个教师,就是第一个进入抽样库的那名教师。在第一个教师被抽出之后,我们再从那名教师所在位置开始,每隔9名抽一个,那么清单上第11、第20、第29和第38名……,共至少200名教师将被陆续抽入抽样库。

在从一份非常长的清单(如市内电话簿)上进行抽样,从前往后计数,选取被调查对象时,如果不是采用一次选取一个的方法,而是采用下例所示的方法进行系统抽样,抽样过程将会变得更快更容易。假定我们要从一本电话簿上抽取有300个住户电话的抽样库。我们还假定电话簿一共有50页,每页有电话号码4列。我们应从每页抽取6(300 ÷ 50 = 6)个电话号码。因为每页都有4列,所以我们可以采用某种任意(即无偏倚)的方案在列内进行抽样。可供选择的方案有无穷多个。例如,我们可以从一页的每一列中,抽正数第10个住宅用户号码和前两列中的倒数第5个用户号码,每页共6个号码,全部50页共300个号码进入抽样库。为了

进一步缩短这令人乏味的抽样过程,我们可以构建一个用于每一页抽样的模版或采用整群抽样,但后者会使抽样误差有所增加(参见 Lepkowski,1988,pp. 76 ~78)。

在使用清单进行抽样的时候,大家必须记住的一个重要的问题是,我们所采用的抽样方法不论是什么,它都必须按列在清单上的全部名单进行。因为一般清单都以姓氏的字母次序排列,这样从清单所列的全部名单中进行抽样,便可选出一个包括全部姓氏的有代表性的抽样库。同样,如果清单以资历的高低排序(资历最长的成员排在清单的前面),那么在用整个清单所列名单进行抽样所产生的有代表性的抽样库,将与成员的资历长短无关。诚如亨利(Henry,1990)所说,研究者在选择抽样间隔时,应该倍加小心,避免使用那种与清单的某种内在的呈周期性特征的性质有一致性的抽样间隔。此外系统抽样还有另外一个优点:当我们用某种研究者感兴趣的特征排序的清单进行系统抽样时,最终得到的样本实际上是一个分层样本,而抽样误差则会因此而自动地有所降低(参见 Henry,1990,p. 101)。

我们可以用反向电话簿对一个比较小的地方性区域(例如邻里地区)进行抽样。对于这样的地区采用 RDD 法抽样,费用往往过高。而用反向电话簿抽样费用则较低,尽管未覆盖误差(因为丢失了未登录的住户)会有所增加,但一般不会超出研究可接受的范围。为了确定每一街区应该抽取的样本数,最好的办法是得到一张可资利用的街道地图。如果地图可显示每一街区的实际住户单位,为了能在最终样本中反映每一街区的住户比例,我们可以采用分层抽样。如果没有现存的可资利用的分层的资料,研究者只需先在地图上计数抽样区中出现的街区数,然后再将计划抽取的抽样库的样本数除以它,得到的商数便是每一街区应抽取的电话号码数。例如,假如某一研究者在一项评估性研究中,有控制的邻里地区和试验的邻里地区各一个。为了在每一地区完成对 300 个调查对象的调查,他必须为每一地区建立一个容量都为 500 的抽样库。如果每一邻里地区共有街区 100 个,那么他便需要从每一街区随机或系统地抽取 5 个电话号码(这一方法假定每一街区的人口数和/或电话数是相等的)。要能恰当地使用一本反向电话簿进

行抽样，我们还需要得到有关每一街区的第一户和最后一户的精确的地址资料。而一张详细到街道层次的地图的优点，正是在于它能为我们提供这样一种我们所需要的信息资料（这样的地图一般可从市政府的计划部门得到）。

现在许多组织机构已经有了他们的雇员或成员的已经计算机化了的清单。在这样的情况下，他们的计算机软件也许可以在整个清单上进行随机或系统抽样，或者也许可以建立一个可用其他软件处理的数据库。此外，某些反向电话簿也存在有某种可资利用的已经计算机化了的形式。

买卖商业性清单

20 世纪 80 年代，买卖清单的活动开始逐渐增长。调查研究者对某些清单尤其感兴趣，这些清单包括城市电话簿、雇主地址清单、用户号码清单和客户名录清单等。为客户建立各种收费的电话调查抽样库，包括 RDD 抽样库的公司也应运而生。此外建立的许多诸如此类的清单也可以按研究者感兴趣的某种特征（如邮政编码）分层。

清单可以购买的确颇具吸引力，但研究者在使用由他人提供的清单时必须十分小心，尤其对清单中可能存在的范围误差要加倍小心。如果研究者选择购买清单作为他们的抽样库，或者作为他们建立抽样库的依据，那么一个小心谨慎的研究者同样也应当是一个小心谨慎的顾客。这就是说他们必须在确定卖方就他们就有关出售的主清单和抽样库提出的各种问题（例如它们究竟是如何建立的？数据库的信息资料是从哪里搜集的？清单中的数据多长时间更新一次？），以及销售方自己承认的清单中的范围误差有多大等给出的答案的确能令他们满意时方可成交。记住，如果与主清单有关的范围误差是与它资料的搜集和编辑方式有关的话，那么以“主清单”为依据进行抽样得到的样本，可能并非是代表目标总体的概率样本。总体上讲，我同意列普科沃斯基（Lepkowski，1988）提出的忠告：“一般来讲，发生研究者对调查操作过程中的这一关键阶段（即建立抽样库）失去直接的控制这样的情况，是不能允许的。”（p. 79）

然而由于商业性清单毕竟给我们提供了很大方便,因此那些使用这种清单的人,不妨可用某种比较谨慎的折中的方法,对购得的清单试作一定的修正,以减少因为未对它作任何修正就直接使用而可能发生的总调查误差。例如我们可以从某一个供应商那里购买一份某一特定抽样区一般总体的指定的电话号码的清单。然后再用某种随机方法取代每个电话号码中的最后两位,或采用某一种加数法,在每一个后置号码的最后两位都加上某一指定的两位数,生成最终进入抽样库的新的电话号码。这种方法不仅比从零开始搜集前置号码和后置号码的信息费用低得多,而且有助于解决因购买的清单固有的未登录号码而产生的未覆盖问题。我们还可以再举一个其他的例子,例如研究者也许可以选购一份登有姓名和电话号码的清单。然后我们可以令调查人员以购得的清单为依据,设法搜寻新的号码,以减少因无法和抽入到样本中的个体取得联系而造成的未联系人数。较之那种不作任何系统地进一步调查就直接使用购买的清单的方法,这种方法可以在一定程度上减少无回答误差发生的可能。

混合式抽样库

正如我们在本书第1章所讲的那样,在过去的十年中,人们对那些将多种抽样和资料搜集方法并在一起进行的调查表现出来的兴趣越来越高。我们吁请人们将不同的调查方式结合在一起使用的主要理由是,各种不同调查方式(个别访谈、电话调查和邮寄式调查)可以彼此取长补短,从而使总调查误差有所降低(参见Dillman Sangster, & Rockwood, 1992; Dillman & Tarnai, 1988)。对那些正在考虑采用什么样的方式建立抽样库的调查研究者来讲,编制用于混合模式的抽样框所需要的时间和资源可能更长更多。例如,对一个边界和前置号码不太一致的特定邻里地区进行邮寄和电话混合式调查,可能就需要采用双重框架法。我们可以用反向电话簿来建立用于电话调查部分的抽样库,当然电话簿上刊登的住户号码只能是它刊印时候登录的电话号码。在建立用于邮寄调

查部分的抽样库时,我们就需要使用市政部门提供的街区/街道地图。这时,我们必须消除那些因为采用两种方法抽样而重复抽到的那些住户。此外,因为电话簿上也刊有抽入抽样库的住户的邮政地址,因此我们可以在进行电话调查之前,给他们邮寄一封信,通知他们已入选为电话调查的调查对象。而在进行邮寄调查时,有时可能需要给调查对象作多次追踪邮寄,这时我们可以在给他们的信件中加上一些项目,询问调查对象家中是否有电话,如果有的话,电话号码是否已经登录?这些信息使调查研究者可以了解用两种不同方法搜集到的数据可能存在的差异。

电话调查者在考虑采用混合调查模式时,应该随时注意日益增长的有关混合式调查方法的研究文献,以了解究竟在什么条件下,“兼顾费用和其他误差时,混合式调查得到的资料才是正当的”(Lepkowski,1988,p.98)。

读者不要把混合调查模式和那种虽然也同样使用邮件,但其目的只是在电话调查进行之前,通过邮件预先通知那些可能的被调查人的电话调查所使用的混合式抽样库混为一谈(这一问题将在第4章作讨论)。例如,如果电话调查的抽样库是以一个专业组织的所有成员的清单为依据建立的,而且该清单上还登有所有成员个人的邮政地址,那么调查研究人员就可能希望用邮件来通知样本库中的人,告诉他们已被选中参加不久将要举行的电话调查。在这样的情况下,将通过邮件事先与被调查人进行联系的方法,与电话访谈结合在一起使用虽然不失为一种明智之举,但这种做法实际上并未建立一个真正的混合式(或双重框架)的抽样库。

抽样库的大小

最理想的电话调查不仅应该是每一个电话号码都可以拨通,而且也应该是每一个选中的调查只需打一次电话就可以联系上且没有一个拒绝我们的调查。当然,对于任何调查来讲,这样一种理想状况无异于天方夜谭。任何一个抽样库,即使是用最近更新的清单,也会有“错”号存在。在抽到的被调查人中,总会有一小部分

在调查实际进行时无法找到。而对于大多数电话调查来讲,要使拒答率为零(甚至低于10%)也同样是不可能的。

由于这些问题的存在,一般抽样库都要比实际希望完成的访谈数大许多。调查研究者所必须承担的责任之一是选择并确定最终样本的样本量。确定的样本量必须足够大,以能对预期的调查目的给予支持(也就是一个抽样误差足够小的样本),但本书的目的并不在于对如何就这一问题做出决策给读者以详尽的建议。如果读者对诸如这样的决策过程不是很熟悉,可以参阅其他那些论述抽样误差和统计功效的教科书。例如在萨德曼(Sudman,1976)撰写的教科书中,有整个一章(章名为“样本量究竟应该是多少”)介绍在设计阶段,如何为那些目的在于估计某些单一的总体参数的调查确定最终样本的容量(在萨德曼书里,样本一词是指最终完成的访谈数,而不是指为了最终完成调查而将要进行处理的抽样库的大小)。

对那些正在或将要进行调查设计,而调查的目的又在于测量双变量和/或多变量之间的关系的研究者来讲,海斯(Hays,1973)和利普塞(Lipsey ,1989)的有关著作为他们诠释了统计效力和样本量的概念。读者可以回忆一下,我们在前面曾经说过,当设计进行一项目的在于估计某些变量在给定的总体中存在水平的调查时,如果我们采用概率抽样,并且所有其他可能的非抽样误差源是恒定不变的,那么调查测量的精确度,将随着样本量的增大而增大。在二分变量(如一个含同意和反对两个答案的有关对堕胎问题的态度的调查题项)的情况下,若简单随机抽样抽出的样本容量为100且被调查人给出的两个答案之比恰好为50/50,那么在置信区间为95%时,抽样误差的范围约为±10%。当样本量为1 000时,抽样误差将降至±3%左右,而当样本量为10 000时,则抽样误差进一步下降到±1%左右。所有这些数字都假设样本抽自一个任意大的总体。如果样本抽取的总体比较小(譬如10 000或更小),随着总体自身大小的减少,某一特定大小的样本的抽样误差也将随之开始减少。

正如我们前面所提到的那样,那些完全采用CATI系统(即带有适当的软件的系统)的调查,也许不必在调查开始之前建立完整

的抽样库。而那些采用 PAPI 法包括许多采用 CATI 法的调查,可能都需要在访谈开始之前就建立能满足整个调查需要的充分大的抽样库。但抽样库中的电话号码的发放,不仅必须以天为单位逐渐发放,而且必须对整个发放过程进行严格的控制(就如在第 3 章和第 6 章中讨论的那样)。

如果我们采用的是这种在调查开始之前就建立起完整的抽样库的方法,那么我们就必须先要能精确地估计抽样库的大小。从实际操作的角度看,我们必须事先产生数量足够的号码,以不必在以后的调查过程中,因抽样库中的电话号码已经放完,但调查预期的样本量仍未完成而一而再、再而三地多次产生电话号码。但我们也不要因此走向另一个极端,生成一个毫无必要的非常大的抽样库。即使我们可以用计算机来生成这样的抽样库,其成本效益也必然是得不偿失的。为了得到合乎实际的调查所需的样本量的估计值,以下问题是我们必须予以考虑的:

1. 计划需要的最终样本的样本量;
2. 样本库中运行号码的击中率;
3. 根据事先确定的被调查人排除的标准,运行中的电话号码最终可能被筛除的比例;
4. 因无回答而被排除在最终样本之外的合格的被调查人人数。

这些信息可以用于下列抽样库容量估计的公式:

抽样库容量估计值 = (FSS)/((HR)(1 - REC)(1 - LE))

最终样本量

最终样本量(FSS)是调查研究者已经确定的达到既定调查目的所需要完成的访谈人数。正如前面介绍的那样,这是一个在一定程度上基于支持数据搜集的资源可利用性所作出的重要决策。

击中率(Hit-Rate)

击中率(HR)是抽样库中可能在运行并可在适当的地方(即是住宅而非营业场所)接通的电话号码的比例的估计值。击中率随抽样区域的变化而变化。在缺乏以前来自某一特定总体的经验可

作参考的情况下,精确地估计 HR 值的方法是从市内电话公司或市内电话簿中搜集广泛全面的信息。例如,用一本反向电话簿,我们可以估计抽样区每一前置号码中运行电话号码的密度。这时,若能对可能未登录的电话号码的比例有一定的了解,也将不无裨益。

他人的经验也可以为我们提供 HR 的可靠的估计值。格罗夫斯和卡汉(Groves & Kahn,1979)在他们进行的全国性的一阶抽样中发现,抽样库中 20% 左右的号码通向住户。NUSL 于 1992 年完成的大规模的全国性 RDD 研究发现,23% 左右的第一阶电话号码是住户号码。后者的 HR 值应该是 0.23。从我自己在美国各地进行的地方性抽样中得到的经验看,HR 值一般在 0.30 ~ 0.50 之间。例如,在芝加哥地区(区码为 312 和 708 的地区),大多数前置号码中的住户号码的密度都很高,有时在一个前置号码的 10 000 个可能的电话号码中,有高达 4 000 ~ 6 000 个号码是住户号码。这时,HR 在 0.40 ~ 0.60。尽管随着与市中心距离的加大,HR 的估计值也可能随之下降,但最好的办法是每遇到一个新的抽样区,都要搜集那里的有关信息资料,并对那里的 HR 值做尽可能精确的估计。在对 HR 值作估计时,还有一个问题也必须清楚,那就是可能的非住户电话号码的比例。在前面有关章节中曾经提到的 1992 年的 RDD 调查中,NUSL 发现,美国住户与非住户电话号码之比为 2:1。

排除调查对象的标准(Respondent Exclusion Criteria)

任何一种排除被调查人,即将某些个人或住户确认为不合格的标准(REC),无疑都将导致抽样库样本量的上升。(这样做并不会产生调查的无回答误差,因为这些个人或住户并非目标总体的组成部分,理解这一点很重要。)被排除的住户或个人的比例通常以 REC 表示,它的计算方法很简单,但有两点必须注意:首先排除的标准,必须根据调查目的预先确定,其次它需要得到抽样地区最近一次普查的资料。例如,如果我们的调查对象只包括 17 岁以上的男性,那么大约所有住户中的三分之一将被排除在调查之外,因为这些住户中没有符合条件的被调查人。这时,REC 的值等于 0.33。许多在都市地区进行的 RDD 老年(59 岁以上)调查发现,在

每5个接通的住户电话中,有4个没有合乎条件的老人,所以其REC值高达0.80。

RDD调查常常被设计用于必须对地理位置进行筛选的某些特定的社区的抽样。这时,调查研究者必须对那些可以通过电话联系,但实际并不包括在抽样区的人的REC值进行估计。

有许多调查,调查的目标并非一般总体中某些特定的部分,而是那些年龄在17岁以上的个人(即所有的成年人)。如果情况确实是这样的话,那么在这样的调查中有效REC值将为或接近零(0.00)。但随着被调查人选择的标准越来越严格,这个值也会越来越高,甚至会逐渐高到我们无法采用从一般总体中随机地抽取电话号码这样的抽样方法。出于各种实际的考虑,我在电话调查中使用的REC值一般在0.10~0.20,在REC值高于(注:原文为低于(below))这一范围时,我则要奉劝那些考虑使用电话调查的人,可能应该考虑采用其他方法进行抽样。当然诸如这样的决策在很大程度上取决于调查的目的和可资利用的信息资源。例如,NUSL最近准备进行一项REC值为0.98的RDD调查,这就是说每50户住户中,将有49户因没有合乎条件的被调查人而被排除。

合格的被调查人的丢失程度

因无回答而导致的合格的被调查人(LE)的丢失程度是在估计抽样库大小时,我们最后一个应该予以考虑的因素。在一般公众中进行的电话调查,即使我们的督导员和调查员都受过严格的培训且经验非常丰富,LE的值也可能主要因为被调查人拒绝回答,而在0.15~0.25之间。经验告诉我们,随着抽样越来集中在中心城市,加之调查访谈的质量不高,LE的值可能达到甚至超过0.50。这时,我们在考虑调查的有效性问题时,首先应当考虑的问题便是无回答误差可能会有多大(参见Groves,1989)。

相反,在进行某些特定总体的电话调查时,LE值可能会很低。例如,在NUSL于1992年进行的一项有关刑事司法从业者的全国性调查中,调查设计时估计的LE值为0.05,而实际调查的结果,LE值为零。这就是说,对抽样库中的每一个人进行的调查都取得

了成功!(关于如何最大限度地减少调查的无回答人数,从而降低LE值的问题,将在第5和第6章讨论。)

估计抽样库样本量

在写本书第一版时,NUSL曾经受客户委托进行一项调查。该调查计划采用RDD法,在芝加哥调查300个非犹太裔和非阿拉伯裔的年龄在25岁及以上的男性白人或黑人。芝加哥的HR值在0.50左右。根据对被调查人筛选的结果,一般一个住户需要联系一次,REC值大约为0.50。最后我们用筛选后的芝加哥市只含男性居民的样本测得有效的LE值约为0.40。我们将这些估计值代入计算抽样库样本量估计值的公式:

抽样库样本量估计值=(300)/((0.50)(1-0.50)(1-0.4))=2 000

得到为完成该次调查要求的300个被调查人的访谈,共需处理的电话号码总数为2 000个。考虑到各种不可预期的偶发因素,我本人则倾向令样本量在这一估计值上再增加10%,建议在如本例这样的调查中,其建立的抽样库应有电话号码2 200个。

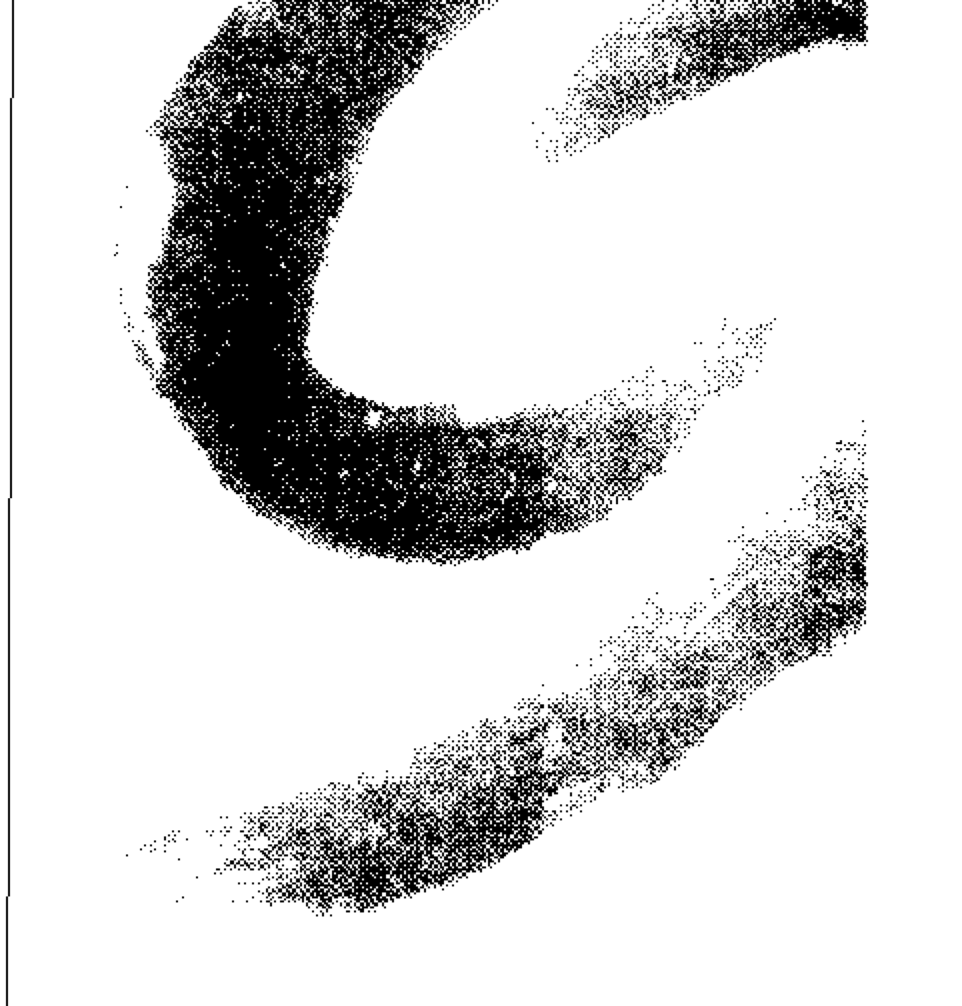
练　习

1)用随机数码表(用人工)建立一个含20个电话号码的抽样库。使用下列前置号码和等比例地分配20个随机后置号:328-、475-、491-和866-。

2)用随机数码表(用人工)建立一个含100个电话号码的抽样库。使用下列前置号码和前导后置数字,再将抽样库按前置号码和括号中列出的比例分层:328-1(10%)、328-3(20%)、475-6(6%)、475-7(5%)、475-8(4%)、475-9(10%)、491-2(22%)、866-0(15%)和866-1(8%)(或用BASIC或其他语言编一个程序建立抽样库)。

3)用市政当局的市内电话簿和最近的有关市政当局的住户电话号码的数据,估计该市政当局的未刊登的电话号码的普遍程度。

4)用习题3)市政当局的有关资料,计算为下面的调查建立的抽样库的样本量的估计值。调查对象是每户中18～29岁的成年人,计划完成访谈的数量为500。假定每5个合格的被调查对象中就有1个拒绝参与调查,而在调查进行期间,每10个合格的被调查对象中就有1个联系不上。



控制电话调查抽样库

Processing Telephone Survey Sampling Pools

抽样库控制问题

用呼叫单控制抽样库

拒访报告表、拒访转变和无回答误差

正确对待在抽样库处理过程中出现的各种结果

抽样库处理结果及调查的回答率和无回答率

混合式调查的抽样库

用 CATI 系统处理抽样库

练　习

我们在第2章给读者介绍了建立电话调查数据库的步骤和方法，而在这一章我们将要给读者介绍如何控制使用电话调查抽样库中的电话号码。对电话调查抽样库的处理过程加以控制的主要目的在于尽可能地降低无回答误差，从而使最终搜集到的数据尽可能地成为抽样框中一个有代表性的样本。读者必须注意的是，未覆盖误差将不受抽样库处理方式的影响，而不是不受代表目标总体的抽样框的方式和/或抽样库从抽样框中建立的方式的影响。

在本书看来，与其他调查方式相比，电话调查的主要优点在于它能使整个数据搜集过程的质量控制集中化。而对调查员和督导人员处理抽样库的过程加以控制，又是这种集中化的质量控制方法的关键之所在。

在第3章我们首先向读者介绍为什么要控制和如何去控制抽样库的问题。其中，我们将介绍呼叫单(call-sheet)，它等同于CATI中的在屏(on-screen)使用方法。呼叫单是一种配置给每一个从抽样库中释放出来的电话号码的表格。在这一节我们将要详细讨论在调查员呼叫每一个电话号码时可能发生的各种情况(处理结果)，以及调查员应该如何掌握和运用这些预先处理结果好的应对各种情况的方法。督导人员(或已经编入CATI系统的程序)正是用这些预先处理结果的方法确定是否(和在何时)令调查员重拨某一电话号码。下一节讨论如何处理由使用拒答报告表(refusal report form)产生的呼叫单。这一内容是比较新的。随后一节深入讨论与抽样库的控制使用有关的几种处理结果的处理方法模式，以及与电话调查的回答率有关的抽样库的控制使用方法。在本章结尾，我想和读者详细探讨一下如何控制使用混合式调查的抽样库，并概括地介绍一下CATI在降低无回答误差方面具有的某些优点。

抽样库控制问题

必须防止调查员随意地从抽样库选取电话号码。因为这样可能会产生相当大的无回答误差。为了达到这一目的，我们必须使用一种正规的方法，系统地控制抽样库中电话号码的提取和使用。

例如,只有借助高度常规化的系统,那些比较难以联系的被调查人才可能得到次数恰当的回呼。如果我们没有这样做,那么在最终的样本构成中,比较容易联系到的被调查人的比例可能会偏高。在一般总体的调查中,比较难联系到的被调查人大多是青年和男性,而比较容易联系到的则是老年和女性。如果我们不能很好地处理这些难以联系到的被调查对象,那么无回答误差便会因为没有联系到原定的某些被调查人而可能有所增加。进而,最终样本的外在有效性(可归纳性)也会受到影响。因为调查关注的专题大多与人口因素(如年龄和性别等)有关,为了减少这样的问题,我们必须在调查中使用正规的系统对调查员的拨号过程进行全程监控。

诸如这样的系统可提高调查员的工作效率,因为它使他们不必再为计划安排重新拨号而分心。这样的系统使调查员再也不必为应该处理哪一个电话号码而操心,从而可以集中精力做好访谈。不仅如此,诸如安排对最初不接受访谈的人进行回访(即给那些最初拒绝接受访谈的人再次打电话,劝说他们接受访谈)这样的工作,也可由控制抽样库的督导人员承担,由他们根据呼叫单上的记录,针对不同情况,一一做出处理。如有必要他们将安排适当的调查员,在适当的时候一一进行回呼。总之,必须由督导人员负责从抽样库中给调查员提供电话号码,以尽可能地避免由调查员自行从抽样库中直接提取电话号码而产生的偏倚。

在进行 PAPI 电话调查时,从抽样库中选号和对已经处理过的电话号码进行分类的工作都是由人工完成的。我们将在这一章对它进行详细的介绍。用人工方法对抽样库的使用进行控制的三个基本步骤是:

1. 在开始进入电话访谈阶段时,应给调查员分配一组数目比较少的电话号码作为他们在最初的访谈时段使用的电话号码;

2. 在每一访谈时段,督导员应该给调查员提供一些备用的电话号码;

3. 在开始下一时段的访谈之前,必须对前一时段处理过的电话号码进行整理分类。经验告诉我们整理分类工作最好由专人负责。

为每一从抽样库中释放的电话号码建立一个单独的呼叫单(如图3.1所示)是各种用于这一层次的控制方法的首选。调查员拨叫的每个号码都有一张单独的清单,这张清单将为调查员拨叫的每一个号码构建一部历史(或称呼叫记录),上面记载着每一个电话号码在每次呼叫时所发生的情况。即使用CATI,系统也会在屏幕上显示一张类似的呼叫单,用以控制和帮助调查员处理抽样库中的电话号码。正如本章后面要指出的那样,CATI和PAPI在调查过程中这一方面的不同之处在于,前者将决策规则编入计算机程序,决策过程由计算机系统自动完成,而后者由人工完成。

读者应当注意,本章所介绍的用以控制抽样库的方法,在某种意义上讲都是人为的选择结果。本人曾经成功地使用了纸笔式的手工分类系统将近15年。其他调查机构,虽然也使用类似系统进行PAPI电话调查,但他们不一定总是使用如此详细的处理结果码,也不一定总是对每个电话号码都穷追不舍(即再处理)。

我们在这里介绍的方法不是一种可供我们使用的惟一的系统方法。在读了本章介绍的内容之后,大家就可以选择设计一些能最好地满足自己需要的经过一定修正的系统。即使使用CATI,为了使电话调查能真正达到高质量,我们也必须对抽样库电话号码的分发过程进行经常和仔细的人工监控。当然与PAPI相比,督导人员在这方面需要付出的时间可能比较少。

用呼叫单控制抽样库

每一个从抽样库分发的电话号码都应该单独打印在一张叫做呼叫单的纸上,调查员在上面记录的有关信息,使督导人员可以据此决定如何处理每一个已经处理过的电话号码。有些调查机构把这种纸片叫做呼叫记录(call record),而另一些机构则把它设计为访谈记录表[interviewer report form (IRF)]。

调查名称

电话号码 问卷#:_____

###-####

试呼次数	日期	时间	处理结果编码	调查员编号
1	__/__	__:__	__ __	__ __
2	__/__	__:__	__ __	__ __
3	__/__	__:__	__ __	__ __
4	__/__	__:__	__ __	__ __
5	__/__	__:__	__ __	__ __
6	__/__	__:__	__ __	__ __
7	__/__	__:__	__ __	__ __
8	__/__	__:__	__ __	__ __
9	__/__	__:__	__ __	__ __
10	__/__	__:__	__ __	__ __

备 注

1. ______________________________
2. ______________________________
3. ______________________________
4. ______________________________
5. ______________________________
6. ______________________________
7. ______________________________
8. ______________________________
9. ______________________________
10. ______________________________

图 3.1 呼叫单实例

读者必须记住"从抽样库分发"的确切含义是,我们并不需要处理抽样库中的所有电话号码。正如我们在第 2 章中讨论的那样,我们提倡在访谈开始之前建立完整的抽样库,除非 CATI 系统可以在访谈过程中自动建立抽样库(但并非所有的 CATI 软件都具有这样的能力)。正因为如此,只有在实际需要的时候,我们才将电话号码从抽样库中取出交给调查员。为了降低无回答误差,进而最大限度地提高外部有效性,我们必须对抽样库分发的所有电话号码都做充分的处理(即进行足够次数的回呼)。假定我们从抽

样库提取的电话号码是所有已经生成的电话号码的一个随机子集，那么，即使为了达到调查规定的访谈完成数量，我们并没有提取抽样库中所有原来已经生成的电话号码，无论是无回答误差，抑或总调查误差，都不会因此而加大。

图3.1是NUSL使用的呼叫单的基本格式的示例。它仿照伊利诺伊州州立大学调查研究实验室使用的表格绘制而成。它与许多调查机构使用的表格颇为相似。其目的在于为调查员提供一种有结构的正规表格，用以记录有关呼叫从抽样库分发的每一个电话号码的重要信息。

记录在如图3.1所示的呼叫单上的内容包括：日期、时间和每次呼叫的处理结果（disposition）。所谓处理结果是指已经编码的每次呼叫结果（即通了但没人接、占线、次序颠倒、拒绝调查和回呼等等）。每次呼叫的调查员的编号（ID number）也应该登录在上面。呼叫单的最后部分作为备注的空白栏，可以记录那些对呼叫单进行分类处理的人，或以后再次呼叫该电话号码的调查员都非常有帮助的信息。

正如我们将在第4章要讨论的那样，呼叫单应与另一张单子，即一张刊印有一段简短的调查导言和一段有关被调查人选择方法的单子配合在一起使用。只有在访谈已经完成的情况下，呼叫单才能被附在问卷纸上。如图3.1所示，电话号码可以打印在单子的左上角，而右上角的空白则可留在以后登记每一份问卷的编号（或个案号）。呼叫单的名称应随调查而异。如果同一组调查员同时进行不同的调查，那么不同调查的呼叫单使用的纸张颜色最好不要相同。

在实际拨叫一个电话号码之前，调查员应该在“试呼次数”列中的某一空白行的适当位置，记录该次呼叫的日期和时间，日期只需记月和日。最理想的情况是在工作室的某一面墙的中央，所有调查员都很容易看见的地方挂一个钟，以便调查员记录调查进行的时间。调查员在拨电话号码之前最后还应该做的一件事，是在呼叫单上填上自己的编号。编号一般是一个两位数，但在调查员人数不多时，只用姓名中的第一个字母代表也可以（在用CATI系统进行调查时，所有这些信息将会自动记录在案）。

记录呼叫的日期和时间是很重要的,因为这将有助于我们随后就如何进一步处理这一呼叫单做出决定。例如,因为在整个调查进行期间,不同次数的试呼之间应有一定的时间间隔。如果每次呼叫的日期和时间都已记录在案,那么督导人员便可以此为根据,决定重新处理某一电话号码的最佳日期和最佳时间。有时,特别是在有些电话可以接通,但总是没有人接听的时候,控制抽样库的人就必须把这张呼叫单暂时搁置起来,以等待适当的日子和时间再作处理。

要调查员登录他们的编号的主要原因在于,它能使督导人员确认是哪一位调查员呼叫了哪一个电话号码,而掌握这样的信息对顺利地解决调查过程中必然出现的那些问题至关重要。在用PAPI系统作电话调查时,这一信息还为我们提供了估计调查员工作效率的依据(CATI系统则为我们提供了计算有关调查员工作效率的各种统计数字的更大可能)。除此之外,由于潜在的心理上的原因,诸如这样的系统将使每个调查员的质量观念得以加强,因为他们知道他们的工作始终处于督导人员的监控之中。

呼叫单上记录的最为重要的信息是标志每次呼叫的具体结果的处理结果码。正是通过对处理结果码的检查,我们才能在每个调查阶段结束之后,迅速而高效地对呼叫单进行分类。在后面的有关章节中,我们将不厌其详地对一组处理结果类别及与之对应的数字码一一给读者进行介绍。对大多数电话调查而言,重要的问题是如何使调查员掌握和使用这些数字编码方案,从而使督导人员能更加专注于对抽样库的控制和管理。而诸如这样颇为正规系统地使用,势必又会使调查员的专业化程度有所加强。

表3.1便是诸如这样的编码方案的示例。读者应该明白,表中所列的与各种处理结果对应的数字编码都是主观决定的。也就是说,这些数字代码本身并无什么特别的含义。

表3.1中列出的一组处理结果码对那些在一般总体中进行的,采用正规的选择方法在每个已经联系到的住户中选择被调查对象的调查来讲是相当典型的。一组类似表3.1所列的处理结果码,既可以用于依据清单建立的抽样库,也可以用于RDD抽样库。根据调查目的的不同,调查研究人员也许需要补充使用一些额外

的处理结果码。例如，在进行同组追踪调查（即一种在一定的时段，如一年之后对同一调查对象再次进行调查）时，对那些因各种原因，在同组调查的前一次调查中曾经通过电话进行过调查，而在再次进行调查时无法再通过电话联系上的人需要给出额外的处理结果码。

就以下表3.1和其他地方所列的处理结果码的含义所作的解释而言，我个人将那个负责控制整个电话号码处理的督导人员看做抽样库的监控者。正如前面所讲的那样，我认为比较好的做法是将包括诸如什么时间把特定的呼叫单发放给调查员、什么时候重新制定处理某一呼叫单的计划，以及什么时候取消对某一呼叫单作进一步的处理这样的事宜交由专人负责。而有关最后一种情况的决策，在一定程度上取决于调查可用于要求调查员因各种原因，如已经接通但无人接听、初次拒访、遇到和比较难打交道的被调查人等而导致的回呼的预算。

表3.1 控制抽样库的处理结果码示例

处理结果代码	代码含义
10	铃响7次后，无人接听
11	忙音，在忙音之后接着呼叫仍然忙音
12	应答机（居住处）
13	住户语言障碍
14	接听者非本户成员
15	住户拒绝；因为接通以后未置一词立即挂断，所以用15H
20	无法接通或为其他运行状态
21	临时性断线
22	营业场所，或其他非居住场所
23	没有合格的被调查人
30	接通而已；需要另约时间回呼，用30A
31	选中的被调查人暂时无法接受访谈
32	选中的被调查人调查期间无法接受访谈
33	选中的被调查人因肢体或精神残疾无法接受访谈
34	选中的被调查人存在语言障碍
35	选中的被调查人拒绝接受访谈
36	不完整访谈；因为被拒而不完整，用36R
37	完成访谈

通了但没人接听

在很多时候,特别是用 RDD 系统进行调查的时候,调查员拨了一些电话号码,却没有进行访谈。因为这些号码中有许多虽然通了,但根本没有人接听。在过去的 10 年中,由 NUSL 负责执行的若干个 RDD 调查中,在调查员拨叫的所有电话号码中,通了但没有人接听的占 1/3 ~2/3。

如表 3.1 所示,调查员在遇到拨通了一个事先规定的电话号码,却无人接听情况时,应该选择代码 10,表示这次呼叫的结果是通了但没人接听。调查员在挂断之前允许电话铃声响多少次取决于调查研究者事先做出的选择。但一般至少应等电话铃声响了 4 或 5 次以后,才允许调查员挂断(参见 Frey,1989,p. 228;Smead & Wilcox,1980),尽管铃声次数的多少与费用无关。在 NUSL,我们要求调查员在铃声响了 7 次之后,仍然无人接听时,才可将处理结果定为通了但无人接听,即选择 10 作为编码。但也会有例外,那就是拨叫那些在以前已经多次拨通,但总是无人接听的电话时,我们要求调查员必须在铃声响了 10 次以上,仍然无人接听时,方可挂断电话。

通了但没有人接的电话在以后的每一调查时段都应重新呼叫一次,一个时段接一个时段地回呼,直至有人接听,或达到事先确定的最多允许可以重新呼叫的次数为止。回呼的次数在一定程度上取决于调查费用的多少和调查持续时间的长短。一般来讲,预算用于回呼的资金越少,通了但从来没人接听的电话号码的比例就越大。对一个调查研究者来讲,困难在于如何在对二者权衡之后做出抉择。因为用作回呼的额外费用的增加并不与回呼次数的回报成正比,相反回呼次数的回报(即接听比例)却会随回呼次数的递增而递减。据格罗夫斯和卡汉(Groves & Kahn,1979)估计,在他们用 RDD 抽样进行的全国规模的方法学测试中,在重拨 12 次以上的电话号码中,只有 5% 的号码是住户号码。在对无人接听的电话号码进行多次呼叫之后遇到的这种住户号码的低击中率,在许多政府资助的大规模 RDD 抽样调查中并不鲜见(Maklan & Waksberg,1988)。而由这种非常难以联系的住户而产生的无回答误差,

将会对总调查误差产生重大影响。一种可供我们选择但并非总是可行或有成效的方法是与电话公司经常保持接触，设法确定那些永远不接听电话的住户的实际状况。对这一问题有兴趣的读者可参阅赛波德(Sebold,1988)的有关著作。他的著作中有关于如何处理复杂的无人接听问题的详细的论述。

忙 音

在多数情况下，忙音是一个好结果，因为它意味着我们最终可以通过电话与某人取得联系。但我们必须提请读者注意的是，只有调查员在听到忙音之后，立即并非常小心地再次呼叫该号码，在听到的仍然是忙音的那些呼叫时，其处理结果的编码才可定为忙音。我们之所以要求调查员这样做，其原因不仅在于它可以令调查员确认第一次所拨的号码是否正确，还在于再次拨号时，存在原来使用电话的人恰好刚刚结束通话的可能。

调查研究人员根据实际情况的需要确定是否需要调查员随后(如30分钟或更长时间之后)再次重拨忙音号。如果我们决定要求调查员随后重拨忙音号，那么我们还必须对允许重拨的次数做出规定。否则的话，呼叫单上处理结果那一列，很快就会被“忙音”码填满。我认为比较恰当的规定是，允许调查员在一个调查段，回呼忙音号1~2次。

有时，调查员有可能会将“快忙音”信号误认为忙音信号，我们应该避免这样的问题发生。快忙音这种音调是某些电话公司用来表示未运行号码，或因为超负荷而引起的线路过于繁忙。如果调查员对这类声调有所警觉，那么就很少会将它们与一般的忙音信号混淆。否则他们就会因为误将快忙音当做忙音而浪费时间进行回呼。一般我们应将快忙音编码定位为未运行号码，而控制抽样库的督导人员将把它从那些需作进一步处理的呼叫单中去除。如果调查在电话线路常常超负荷的地方进行，那么我们便应该对它另作处理。这时我们应该给快忙音另外一类编码，以便我们在另一调查段对以往听到的快忙音信号加以确认。

应答机

由于电话应答机的价格迅速跌落，致使使用它的公众越来越多。1992年，据一些市场研究估计，一半左右的美国住户使用应答机。此后，这一数字逐年递增。我本人进行的有关应答机的调查发现，最有可能拥有（和使用）应答机的是30岁以下的成年人，他们大多是白人且有比较高的收入和文化程度。这些有应答机的人中的大多数（70%～75%）之所以用应答机是为了不错过来电（这些人乐于与人交往）。与之相反，其中的少数人（25%～30%）则声称，他们使用应答机的主要目的是筛除不想接听的电话（这些人乐于独处）。此外，应答机主似乎一般与未完成率无关，通常回呼10次，或更多次便可完成对他们的调查。

如果应答机上的留言告诉我们，电话的那一方是一个营业场所，而规定的调查对象却是住户，这时，该次呼叫的处理结果码就应编为非住户，而不是应答机（见表3.1）。在调查员根据应答机给出的留言不能确定对方的确切身份，或仍然有所怀疑，或可以确定对方是一个住户时，应当将编码选为应答机。这一类编码给调查的监控者传递的信息是：这个电话号码最终将可能会有一个访谈完成。

调查员在遇到应答机时，他必须立即决定是否要在应答机上留言。不言而喻，留言的主要目的是为了与对方预约下次通话的时间。迄今为止，没有任何研究可以确切地告诉我们是否应该在应答机上留言。我个人比较倾向要求调查员在应答机上留言，因为这样可以使调查员与被调查人之间的联系具有某种个人交往的性质，但留言应该简单、有礼貌和标准化。换一句话说，我们应该指导调查员在遇到住户家中有应答机时如何留言。留言应当与调查的引导性用语即所谓的开场白类似。它应当明确说明调查的执行机构、调查的目的和重要性，并提醒对方他们将会再次来电话。

调查的监控者必须特别关注那些接通到应答机上的电话号码。通过考察呼叫单上与试叫次数和处理结果码对应的具体呼叫时间（几点钟）和/或日期（星期几），我们可以断定一个电话号码是否总是由应答机来接听。如果确实如此，我们便可以把这个电话

号码暂时搁置一边，将它安排在一星期中以前没有呼叫过的那一天或一天中以前没有呼叫过的那一时间再进行处理。当然，最终是否可以这样做，在一定程度上取决于调查访谈实际可延续的时间。

住户语言障碍

一般来讲，调查的抽样区域越接近市中心，调查员就越有可能遇到住户的电话由一个不能讲英语或无法用英语正确地回答问题的人接听的情况。

我们有理由认为，在某些场合，某些被调查人会以“不会讲英语”做借口，躲避调查员的访谈。如果这样的情况发生了，那么这次呼叫实际上遇到了一次拒访，因为被调查人并非真的不会说英语，而是以它作借口。遗憾的是，至今我们还没有什么标准化的方法，可以帮助调查员识别这样的情况。

当语言真的成为一种障碍的时候，调查员就应当将该次呼叫的处理结果码定为语言障碍(13)。即使调查员对当事人是否诚实有怀疑，他也不应当和他们发生争辩。如果经费允许，调查监控人员应该安排在另一调查时段，再一次拨叫这一电话号码，也许碰巧在第二次呼叫时，能够遇上一个可以(且将)用英语交谈的人。这一方法在使用系统户内选样方法(我们将在第4章对它进行讨论)时特别有用，因为最终被选作调查对象的那个人也许是会说英语的。

有些调查，除了雇用讲英语的调查员之外，还雇用了讲西班牙语的调查员，例如NUSL在芝加哥进行过的调查，因为芝加哥有为数众多但身份多少有些未经正式证明的西班牙裔人口。NUSL的经验表明，我们完全可以用英语对半数以上自称为西班牙人的被调查人进行访谈。因此我们有理由认为，在一个一般人口总体的调查中，即使我们没有条件进行双语调查，由此而产生的非英语住户的丢失问题也不会很大(非英语访谈要求我们将问卷翻译并打印成第二种语言。那些可讲两种语言的调查员，不可自行将问卷从英语翻译成另一种语言。如果调查员自行其是地临场翻译问卷，原有问卷的标准格式就会丢失。这样，不同的双语调查员使用

的问卷的措词和格式就会各不相同,非英语的被调查人听到的问卷版本也会因此而各不相同)。

接听者非本户成员

偶尔也会发生接听电话者并非本户成员这样的情况,调查员被告知本户的主人目前没有在家。这种情况多发生在住户的主人外出,当时家中只有应主人之请到来的亲朋好友,或应主人之聘临时在家照看小孩的保姆、钟点工、清洁工等人时。在遇到这样的情况时,调查员应说明致电的原因,并设法弄清主人返回的确切时间。这些信息应该记录在呼叫单上的备注部分。此后,假如实际调查仍在继续的话,监控者可以酌情要求调查员在适当的时候进行回呼。

住户拒访

在一般公众的电话调查中,大多数拒访都发生在电话刚刚接通之后但在户内指定的被调查人选出之前(参见 Collins, Sykes, Wilson, & Blackshaw, 1988)。因此,我们将这种情况的处理结果码定为住户拒访。研究表明,一般在这种类型的拒访中,初次拒访的比例在 2/3 ~ 3/4(Lavrakas, Merkle, & Bauman, 1992)。在遇到这样的情况时,调查员应该在呼叫单上写明拒访的性质(正如我们将在本章的后面部分所要介绍的那样,调查研究者也许可以选用一种被称为拒访报告表的表格,让访谈员在上面记录他们遇到的每一次拒访的具体情况)。人们经常会这样说:“我们不感兴趣”,“我们没有时间”或“我们不接受调查”这样的话。有时甚至任何话都不说就挂上电话。这四种情况在拒访中所占的比例分别为 40%、20%、10%和 10%(Lavrakas 等, 1992)。在直接挂上电话这种拒访中,其中有一定比例是因为非英语住户所致。因此我们必须令调查员区别不同的情况,给不同类型的拒访以不同的处理结果码(如表 3.1 中的编码 15H)。这就使监控者能在重新分配这些呼叫单时有所选择,如果需要,便可考虑安排双语调查员做回呼。当然,选择的前提是我们的确有这样的调查员。

从样本分析处理的角度看,调查研究者事先就应该将试图进

行的访谈归为拒访的标准做出规定。根据以往的经验，不管造成拒访的原因是什么，其中的大多数似乎都与致电的时间不合适有关。调查员也许在一个对方不方便的时间给他们打电话和/或接电话的人恰好又是一个不愿合作的人。

正如我们在第 5 章将要讨论的那样，我们必须对调查员进行培训，使他们学会在遇到拒访时，如何在最终放弃某次实际的呼叫之前作“礼貌的坚持”。一旦调查员真的觉得某次呼叫已无法进行时，他/她就可以中止交谈，不过在终止之前应该做如下的表示：“好吧，对不起，这一次打扰了。”

对那些调查时间延续比较长的调查（2 或 3 个星期），许多调查机构都有一种处理住户拒访的通用法则，那就是在大约一星期之后再做第二次呼叫，这时电话有可能被愿意合作的人接听。在遇到强烈（但不是固执的）拒访时，最好将这些人交由那些在寻求被调查人的合作方面有特殊诀窍的调查员来处理。NUSL 实际采用的办法是，凡是调查员被住户明确告知“……以后请不要再来电话”的呼叫都不包括在拒访转换试呼中，而是把它们记录在拒访报告表（refusal report form）上。正如我们下面要讨论的那样，拒访报告表（RRF）对那些以后回呼这些号码的调查员特别有帮助。不论我们以后做出的决定是什么，在处理抽样库的过程中，我们都应对住户拒访问题予以特别的关注，因为拒访率必定会对潜在的无回答误差产生影响，进而危及调查的外部有效性。

未运行号码

识别未运行号码的难易程度，取决于在抽样区内运行的电话公司实际采用的表示未运行号码的方式。在比较接近中心城市的地区，大多数电话公司都用各种录音信息来通知来电者“拨叫的号码已经停止使用”。有时，有的电话公司也会采用一些其他作用类似的方法。也有其他一些公司，用一种异乎寻常的噪音来表示已经停止运行的电话号码。而在抽样区域位于农村地区时，未运行号码也许没有什么特别的表示，只是铃声一直在响，却始终没有人接听而已。未运行号码是居电话铃声响了却始终无人接听之后，一般公众的 RDD 调查中各种处理结果码中出现频率第二高的代

码。每当一个调查员在拨叫一个电话号码,而它并非一个正在运行的号码时,他便应当将未运行处理结果码记录在该号码的呼叫单上。此外,这一编码也应包括那些用录音通知调查员,原号码已经更改了(且已经给出了新的电话号码)的那些电话号码。在大多数 RDD 调查中,调查员不应当继续呼叫那些已经更改过的电话号码,因为这样做有可能会以某些未知的方式使抽样概率发生变化。

在多数情况下,未运行号码一经发现,督导控制人员便应立即将这一呼叫单从一组需要做进一步处理的呼叫单中取出。尽管存在第一次呼叫拨错了号,或电话公司随后又将原来未运行的号码发放给了另一个住户的可能,一般来讲,在抽样库容量足够大的 RDD 调查中,我们几乎没有必要对任何未运行的号码进行回呼。这是因为拨错号和重新发放未运行号发生的可能性是微乎其微的,加之这样的情况发生都是随机的,所以即使这样的情况发生了,也不会对最终样本中的范围偏倚有什么影响。更何况从经济角度上看,重拨所有未运行的号码也是得不偿失的。

临时停机

若录音信息告诉调查员电话号码临时停机或"正在检修",这时我们应该使用不同于未运行处理结果码的另外一种单独一类的处理结果码。如果调查延续的时间长到足以使我们有机会再作一次努力的话,督导控制人员应在若干天之后,将这些呼叫单作二次分配。

非住户

在一般公众的调查中,调查员必定会遇到接电话的那一方是一个营业场所或其他非居民单位(如医院、图书馆和政府部门等)。依据抽样区域的不同,运行中的电话号码中非住户的电话号码所占的比例也会有所不同,一般占 1/4 ~ 1/3。所以,这一类处理结果码在使用 RDD 抽样库的电话调查中时,出现的频率一般都比较高。调查员应该了解如何去解释,为什么有关一般公众的调查不能在非住户电话上进行:因为它违反了公民调查自身隐含的概率抽样的假定。每当应答机中的录音告诉我们电话打进了一个营业

场所和其他非居住场所时，呼叫单上便应将该次呼叫的处理结果码定为非住户。如果我们听到从电话那一头传来了调制解调器或传真机发出的尖叫声或嗞嗞声时，一般就可以确定，在呼叫单上该次呼叫的处理结果码应该为非住户，尽管有时有的住户也许也会安装这类机器，但一般也都是出于商业目的。

由于转呼技术(call forwarding technology)的使用，使住户和非住户的区别变得复杂起来，因为电话上编好的程序可以令打到某一电话号码上的电话，自动转到另一个号码上。在 1992 年，全美每 20 个住户中，就有 1 户登记使用这项服务。例如在 RDD 调查中，我们可能拨叫的是一个住户的电话，但在通了以后，却自动转到了一部营业场所的电话上，或者情况恰好相反，拨的是营业场所的电话，却转到了某一住户的电话上。在这两种情况中，每一种情况的处理结果码的类别都应当取决于直接拨叫的那个电话而非转接的那个电话的具体性质。如果一个营业场所的电话被转移至一个住户，那么这次呼叫的处理结果码就应该是非住户，访谈应该就此终止。如果住户号码转移到了某一营业场所，那么访谈可以继续进行，不过，如果可能，最好还是在以后另择时间再作访谈。正如第 4 章将要讨论的那样，一个警觉性较高的调查员可能会把区别这些差异作为例行工作的一部分，对他们自己呼叫的电话号码一一进行校验。一般登录有非住户处理结果码的那些呼叫单，不需要在以后再作进一步的处理。

无合格的被调查人

根据一个电话调查被调查人选择的特定要求(参见第 4 章)，有时电话可能会打进一户没有任何一个合格的被调查人的住户。这一结果应该记录在呼叫单上，监控者应当取消对这一号码作进一步处理。

在对一般的人口总体中年龄在 17 岁以上的成年人作调查时，调查员便会偶尔将电话打到一户说户中没有一个年龄在 17 岁以上的成年人的住户中。因为这些情况中的大多数，与调查员通话的都是孩子，所以尽管我们已将这些呼叫单的处理结果码编为了“无合格的被调查人”，但为谨慎起见，监控者应在后面的调查时

段,再安排一次回呼。在回呼再一次确认联系的住户中的确“无合格的被调查人”之后,监控者方可最后决定终止对这一呼叫单作进一步处理。

接通而已

对大多数调查而言,在所有完成的访谈中,只有不到一半的访谈是只进行了一次呼叫后就全部完成的。而从发展趋势看,这一比例虽然不是很快,但似乎在逐年下降。每当我们打通某一住户的电话,并宣读了标准的有关调查的引导性开场白,以及挑选调查规定的合格的被调查人的步骤之后,却被告知指定的被调查人目前无法接受访谈时,呼叫单上的处理结果码就应定为“接通而已”。这就是说,这一结果并没有作为拒访处理。在这些场合,关键的问题是调查员应该试着确定与被调查人联系的最佳时机,并将这一信息记录在呼叫单上。例如这些信息可能是“晚上 9 点之后”或“只可以在周末”。这些信息可以为监控者在考虑何时再次处理这些呼叫单时提供帮助。此外,调查员应该尽可能地记录一些可识别被调查人身份的信息,如性别、姓名等,为随后再次对这一呼叫单作处理的调查员提供一些帮助。

有时,调查员可以确定回呼的确切时间。如果情况果真如此,对这样的结果采用一种独特的编码(如表 3.1 中的 30A)常常会对监控者确认这些呼叫单的具体结果,并保证对它们在最正确的时间再次进行处理有所帮助。

临时不在

应该对在调查员打电话时(如上面所讨论的那样)临时没有在的被调查人使用单独一类的处理结果码,以区别于那些几天或更长的时间都不在的被调查人。临时不在的原因可能是患小病、短假或出差等。如果调查时间可以延续至这些被调查人回来之后,监控者就可以将这些呼叫单暂时搁置,以等待在合适的时候再对它们进行处理。

长时间不在

与临时不在不同,有时被调查人在调查结束之前总是不会在(如长假、在国外出差等),这时调查员应将这一情况的处理结果码编为长时间不在,而监控者将不再对这些呼叫单作进一步处理,除非原来计划的调查时间延长了。

残疾障碍

有时一个人虽然在人口学意义上(如年龄、性别等)符合作为被调查人的要求,但因为在生理或精神上患有某种疾病而无法参与实际调查。在遇到这样的情况时,我们便应使用残疾障碍这类处理结果码。一般来讲,这类处理结果码也应包括那些住院时间超过调查持续时间的那些被调查人。凡遇到这类结果码时,监控者一般都应当停止对呼叫单作进一步处理。当然调查研究者事先必须就遇到这样的情况时,是否应在同一户内寻找另外一个合格的人作为被调查人这一问题做出决定。而这种决定,必须与整个抽样设计所持的原则相一致。

不过读者也应该注意到,有时候没有能力参与调查是临时性的(如被调查人喝醉了)。在遇到这种情况时,调查员不应该再继续进行调查,但必须在呼叫单上注明被调查人的残疾障碍只是暂时的。监控者应在以后某个适当的时间,对这些呼叫单再作处理。

被调查人语言障碍

正如我们在前面指出的那样,单独一类处理结果码被分配给了那些因语言问题而无法确定户内是否有合格的说英语的被调查人的呼叫单。而在另外一些情况下,调查员虽然可以顺利地确认户内按规定应选择的那个被调查人,他所了解的情况只是那个被选出的调查人不会说英语。例如,一个说英语的年轻人(他和父母生活在一起)也许接了打进去的电话,并帮助我们完成了被调查人的挑选。但按程序选出的被调查人可能是年轻人的父亲,而他的英语也许不够好,以致无法实际参与访谈。在遇到这样的结果时,我们应将处理结果码选为语言障碍。对记录有这样信息的呼叫

单,监控者不应再安排作进一步的处理。

在遇到有残疾障碍这样的处理结果时,调查研究者不仅应事先就是否在该户选择另一位说英语的成员作为被调查人这一问题做出决定,而且做出的决定还必须与抽样设计所持的原则一致。例如,如果调查的目标总体是抽样区域内的所有说英语的居民,那么调查员在执行被调查人选择标准时,就只应该考虑那些说英语的人(参见第4章)。在本例中,虽然我们首先选中的是一个不会说英语的成年人(实际上是一次误选),但因为调查的目标总体仅限于那些说英语的成年人,所以我们并未对他进行调查,这时调查的范围误差并不会因此而增加。然而,如果我们没有给户中另一个说英语的成员提供入选样本的机会,那么范围误差就会因此而增加。

选中的被调查人拒访

在那些用特定的程序从户内选取特定的被调查人的调查中,有时指定的被调查人会拒绝参与调查。这样的情况约占RDD调查中最初拒访的1/3。在出现这样的结果时,呼叫单上的处理结果码应定为被调查人拒访。这一结果不同于在户内规定的被调查人确定之前发生的住户层次的拒访。

对于监控者来讲,重要的问题是区别住户拒访和被调查人拒访这二者。后一种拒访也许很难转变,因为它是指定的被调查人本人给出的,而非该户接电话的人给出的。就所有类型的拒访而言,我们都应该要求调查员对它们的性质加以简短的说明,如果调查使用了拒访报告表,那么我们就应该要求调查员将拒访的情况填在拒访报告表上(此表将在后文讨论)。这一信息将对监控者考虑如何选择最佳(如果确有最佳的话)时机再次对这一呼叫单作进一步处理,以及另一位访谈者在随后的呼叫时,更有成效地为转变拒访者的态度所作的努力非常有帮助。

不完整访谈

访谈一旦开始,一个优秀的调查员很可能成功地说服一个本来不太愿意接受访谈的被调查者最终完成访谈。然而,有时被调

查人可能无法或不愿继续从头到尾回答问卷上列出的所有问题。

在第一种情形中，可能在调查进行的过程中发生了一些什么事情，例如婴儿大哭不止，致使调查无法继续。如果调查员处理得当，被调查人通常都愿意另择时间完成调查。这种情况应该记作不完整的调查，而且调查员应该在呼叫单上注明下次什么时候再次呼叫比较合适。监控者应在随后为原调查员确定进行回呼的恰当时间。当然，其前提是回呼的确是可能的。

在遇到第二种情况时，被调查人不仅拒绝继续进行调查，而且也拒绝另行安排时间继续调查。这时调查员也应将这一结果的编码定为不完整访谈，并且应该为它选择一种单独记号（如表 3.1 所示的 36R）表明这类不完整拒访是一次真正的拒访。在这样的情况下，如果调查使用拒访报告表的话，那么调查员便应当将这一情况填写在拒访报告表中。经验告诉我们，这些类型的不完整访谈一般都发生在老年人身上，因为他们有时会在调查过程中逐渐感到疲劳和/或变得烦躁不安。令我们感到安慰的是，各种类型的不完整访谈并不经常发生在那些训练有素的调查员身上。在不完整访谈确实发生时，监控者也许应该与原调查员就发生的情况作一些讨论。比较明智的做法是在审阅呼叫单和拒访报告表（如果拒访报告表存在的话）之后，再安排进行一次回呼。

某些未完成访谈的发生往往是因为被调查人和调查员之间的个性有冲突。例如，老年妇女在遇到年轻的男性调查员时，常常会感到紧张。在这样的情况下，监控者应当安排另外一个调查员作重访，再次尝试与那个曾经拒访的被调查人进行交谈。

调查完成

在 PAPI 式的电话调查中，一旦一个调查员完成了呼叫单规定的调查之后，便应在呼叫单上填入专门表示调查完成的处理结果码，然后再将它附在完成的问卷上。此后，调查员应对这一份问卷进行编辑整理，对问卷上所有开放式问题的答案的编码一一予以确认。在完成所有工作之后，调查员应将问卷识别号（即编号）填在呼叫单上。注意，识别码必须是惟一的，不可重复。例如，在图 3.1 所示的呼叫单上，右上角的空白处就是专门用来填写问卷编号

的。

在再次呼叫附在一份已经完成调查的问卷上的呼叫单上的电话号码时(即对已经完成的访谈进行校验,这一问题将在第6章进行讨论),偶尔会遇到接电话的人说,他们从来没有接受过任何调查。这种情况的发生,固然有可能是因为住户不愿和我们合作进行调查的核对而故意这样说的,但也不排除它是因为在PAPI调查中的一部分调查员的漫不经心而导致了呼叫单的张冠李戴,即没有将填好的呼叫单附在它应该附的那一份已经完成的问卷上。正如我们在第5章要讨论的那样,当调查员在一个一切都井井有条的工作场所工作时,诸如这样的问题便会大大减少。

简化呼叫单

有许多调查,特别是那些由学生执行的调查,使用如图3.1所示的专业呼叫单也许过于繁琐。如果一组调查员在相当短的实地调查时间内(如一个星期或更短的时间内)只做一项电话调查,那么我们就不必多此一举,因为使用这样一组抽象的数字处理结果码未免有些小题大做。

当然,这并不是说我们不需要使用任何呼叫单,而是说我们只要使用如图3.2所示的比较简单的呼叫单就足够了。在使用简化呼叫单时,临时调查员只需做一些比较简单的记录,向监控者说明每次呼叫发生的情况就可以了。与确定每一次呼叫结果的编码不同,他们只要写下诸如"无人接听"、"忙音"、"停机"和"拒访"等这样的简单信息就可以了,当然这会使抽样库控制者(一般都是教授)在对呼叫单进行整理分类时花费更多的时间,并且丢失诸如每次呼叫的确切日期和时间这样的历史信息,但对一个只需对处理结果做一些一般分类的调查而言,这种简化呼叫单所能提供的信息已经足够了。

调查名称

电话号码 问卷#:______

###-####

调查简介和被调查人选择方法打印处

试呼次数	具体处理结果	调查员编号
1	______	______
2	______	______
3	______	______
4	______	______

图 3.2 简化呼叫单示例

小 结

比较理想的情况是，在实际调查结束时监控者只从抽样库中取出了数量最少的电话号码，便完成了设计规定的应完成的调查人数。监控者一旦在呼叫单上发现了诸如未运行、营业场所、完成和语言障碍这样的处理结果码，便应立即终止对这些呼叫单的处理。监控者应分期分批，并限量地调用和分配新的呼叫单。调用新呼叫单的目的固然是为了使调查员的工作效率能达到我们规定的水平，从而使我们能在规定的时间完成整个数据搜集工作，但调用必须以使无法联系的被调查人（即在样本容量已经达到规定的数量时那些仍然无法完成访谈的人）的人数最小的方式进行。

那些已经联系过但始终未能完成对指定的被调查人的调查，以及那些电话可以接通但始终没有人接听的呼叫都有可能使无回答误差有所增加。在大多数调查中，那种认为应使所有呼叫单的最终处理结果都能为完成调查，或其他具有终结含义的处理结果的想法是不切实际的。因此，监控者的责任是在实际调查结束时，使仍待处理的呼叫单的数量保持在最低限度。在调查结束时发生的问题仍未得到解决的呼叫单数目的多少，在一定程度上取决于预算限定的回呼次数。而对一个监控者来说，他的责任是使来自

抽样库的每一个电话号码的拨叫次数,能达到呼叫单上规定的允许拨叫的次数。

最后我想说的是,每一个电话号码使用一张单独的呼叫单是减少潜在的无回答误差的必要条件,但并非充分条件。对于抽样库控制所给予的专业关注越多,电话调查过程中的各个方面达到的质量就可能越高。

拒访报告表、拒访转变和无回答误差

总的来讲,与个别调查一样,大部分电话调查的无回答是因为拒访,而非无法联系(Groves & Lyberg,1988)。此外,在过去的20年中,无回答率虽然上升幅度很小,却在持续不断地上升。有鉴于此,在过去的20年间,人们也在始终不断地设计一些新的方法,尽可能地减少因拒访而引起的各种问题(参见Lyberg & Dean,1992)。使用某种结构性表格便是其中一种。通常我们把这种调查员在遇到拒访时填写的表格叫做拒访报告表(RRF)。这种表格不仅可以为监控者和调查员在以后为转变拒访而作的努力提供帮助,也可以为调查研究人员更多地了解掌握潜在的无回答误差的性质提供帮助。

格罗夫斯(Groves,1989,p. 218)的观点虽然有悖于直觉,但无疑是正确的。他警告我们:通过转变拒访(对未曾联系到被调查人作的额外的回呼)来减少无回答人数,可能会因为无回答误差和其他调查误差(如测量误差)的增加而适得其反。例如,从转变后的被调查人那里搜集到的数据质量可能低于其他被调查人。而调查测量误差的增加则有可能抵消回答率增加所产生的效应(参见Blair & Chun,1992)。遗憾的是,我们难以确定这样的事究竟会在什么时候发生。有关这些问题的详尽讨论不仅十分复杂,而且也超出了本书讲授的范围,水平高的读者可以参阅格罗夫斯的有关著作(Groves,1989,pp. 133~238)。

如果某一调查研究者希望把RRF引进抽样过程,那么我想告诉他的是,究竟应该在表格上记录一些什么问题至今仍无明确答案。也就是说,时至1992年,在有关调查方法的文献中,有关RRF的介绍

仍是凤毛麟角。在了解这一事实之后，我想奉劝那些对此感兴趣的读者应该关注下面的讨论，并把目光转向未来的有关文献。

图3.3是NUSL使用RRF的实例。一个如图3.3所示的表格既可用于PAPI调查，也可用于CATI调查。RRF应由调查员在遇到拒访（或部分拒访）之后立即填写。在使用RRF时，首先应该填明户内有什么人，什么人拒绝进行调查。但这一问题并非总是很清楚的，而且在很大程度上取决于调查员在终止通话以前已经得到的信息。如果调查员在被拒绝时尚未完成对被调查人的选择（参见第4章），那么应在RRF上写明，该次拒访属于住户拒访。（如遇到下面这样的情况，则应视作例外：假如某一调查规定的被调查对象是每一住户内的任何成年人，在这样的情况下，调查员在与某一户的联系过程中，即使还未执行正式的被调查人的选择程序，但已经知道那个拒绝访谈的人是该户内惟一的成年人时，这样的拒访便不应再记作住户拒访，而应记作被调查人拒访。）

调查员也许会给出某些有关拒绝调查的被调查人的人口学特征的编码，但只有在调查员在一定程度上真正掌握了有关这样的信息时才可以这样做。已经作过的初步研究告诉我们，在大多数情况下，调查员可以精确地给出拒访者的性别、年龄和种族的编码（Bauman, Merkle, & Lavrakas, 1992; Lavrakas 等, 1992）。当然在我们做出这种方法是否真正有效的最后判断之前，还需要作更多的研究。如果这些人口学信息的确是比较精确的，那么监控者便可据此决定由什么调查员来试呼什么样的拒访者，以促使他们转变态度接受调查；而调查研究者则可据此探讨拒访可能产生的无回答误差。例如，初步的研究表明，与最初拒访者同一种族的调查员更容易促使拒访者转变他们的态度。不仅如此，如果所掌握的拒访者的人口学特征与调查测量的特征有关，那么调查研究者还可以通过考察未转变的拒访者的人口学特征，来探讨无回答误差对调查质量可能产生的影响。然而，读者必须了解，在使这些策略和方法能真正有效地为我们所用之前，需要我们了解和掌握的东西还很多。

除上面这些有关拒访者的人口学信息之外，调查员也可以在拒访报告表上填写一些其他的有关信息，如拒绝的强度（激烈程度，见图3.3）、文字评语和对其他一些问题的答案。这些问题主要

为了帮助说明在电话挂断之前双方交谈的具体情况（如果曾经有过交谈）。调查员也可以在上面就采取什么办法来实现拒访者态度的转变提出自己的建议。

调查名称

调查员 #：______________

1. 拒绝接受调查的人是否过了最后一次（最近一次）生日？

是	……35
没有/不能确定	……15

2. 拒绝者的人口学特征：

性别		年龄		种族	
女	……1	儿童	……0	亚裔	……1
男	……2	成人<30	……1	黑人	……2
不能确定	……9	30～59岁	……2	西班牙	……3
		60岁或以上	……3	白人	……4
		不能确定	……9	不能确定	……9

3. 拒访原因：____

____

4. 拒访强烈程度：	很弱	1	2	3	4	5	6	7	很强烈
被调查人态度：	很有礼貌	1	2	3	4	5	6	7	很粗鲁
	没有生气	1	2	3	4	5	6	7	很生气

5. 你是否告诉被调查人：

		是	否
A.	他或她是怎样被抽到的？	……1	……2
B.	在例行的介绍之外介绍了调查性质和目的？	……1	……2
C.	为被调查者保密？	……1	……2
D.	调查数据如何使用？	……1	……2
E.	举证调查的组织者/资助者？	……1	……2

6. 为了寻求被调查人/住户的合作，如果我们准备再与拒访者交谈一次，你有什么建议？

____

____

____

（务请将此表附在呼叫单上，并交给督导人员）

图 3.3 拒访报告表（RRF）示例

调查研究者应该事先对以下两个问题做出明确规定。这两个问题是应该对什么样的初次拒访者进行回访和选择什么样的调查员进行回访,使他们转变态度接受我们的调查。例如NUSL的有关规定中规定,在可通过回呼进行转变工作的拒访户中,不包括那些明确告诉调查员"别再给我们打电话",或用某种其他方式明确地表示以后不希望再次被打扰的这样的拒访户。NUSL的研究人员认为,拒访者的这一要求应该受到尊重,但对除此之外的初次拒访者都应该尽力促使其转变。不过我们认为监控人员应该将那些怀有敌意的非常粗鲁的初次拒访者排除在拒访转变者之外。那些非常强烈地表示拒绝接受调查的被调查者占所有初次拒访者的10% ~15%。请注意,格罗夫斯和莱伯格(Groves & Lyberg,1988)认为应该在从事调查方法研究的研究人员中,更多地开展有关为实现转变初次拒访者的态度所做的各种努力的道德问题的讨论。

在确定由什么调查员进行拒访者的转变工作时,有几个问题应当考虑。首先那些比较能干的调查员往往把这样的工作看做一种挑战,非常乐于去做。其次,如前所述,在安排某一特定的转换工作时,调查员本人的性别、年龄和种族应在考虑之列。此外,发生在白天的初次拒访,往往也许比较容易转变。同样,在初次拒访和最终转变之间重新联系次数较多,以及那些初次拒访的原因并非对调查的领域缺乏兴趣的拒访也比较容易转变(Lavrakas 等,1992)。但读者必须注意,我在这里谈的几点,都只是一些建议性的经验法则,若要将它们应用于实际调查,还有待进一步试验和证明。迄今为止,尚无有关拒访转变的成功率的确切数据,尽管格罗夫斯和莱伯格(Groves & Lyberg,1988)认为它在25% ~40%,而NUSL的这一数字则在25% ~35%之间。在马里兰大学的布赖和程(Blair & Chun,1992)的报告中,这一数字为25% ~35%。弗瑞(Frey,1989)则认为这一数字应为40%左右,他与亚历山大、西波德和普法夫(Alexander,Sebold & Pfaff ,1986)报告的数字相似。而科林斯等人(Collins 等,1988)则认为,在英国这一数字不足20%。

只有在究竟将资源用于拒访者的转变以降低潜在的无回答误差(但有可能导致其他的误差的上升),还是将它用于其他投资(如提高调查员的酬金,聘用更好的调查员)以降低那些因减少无回答

误差而有可能上升的其他误差这二者之间进行权衡利弊之后,方可决定是否要进行拒访者的转变工作。

正确对待在抽样库处理过程中出现的各种结果

对于调查研究者、督导人员和调查员来讲,如果能对出现在呼叫单上的各种各样的处理结果的相对比例抱有比较切合实际的希望,无疑将会对他们的工作有很大帮助。否则他们很容易因为发现许多拨叫并未像他们希望的那样,导致访谈的最终未完成而感到沮丧。不仅如此,抽样库监控者还应该学会防患于未然,在那些可能导致无回答误差和/或范围误差可能增加的问题发生之前,设法找到并解决它们。

呼叫单处理结果的特定分布在很大程度上取决于在联系合格的被调查人过程中抽样库的效率。例如,根据抽样区域内与每一前置号和在每一前置号内运行的后置号的号码范围相关的前置号运行号码比例分层的 RDD 抽样库(正如第 2 章所讨论的那样),所含的未运行号码数量远远低于简单(未分层)的 RDD 抽样库。米托夫斯基-瓦克斯伯格(Mitofsky-Waksberg)的二阶 RDD 抽样库也具有与此相同的优点。一个根据一份清单或一本电话簿建立的抽样库电话号码的处理结果,可能会因 RDD 抽样的结果而异。不仅如此,呼叫单处理结果的具体分布也会因调查的地理区域不同而不同。

完成调查规定的样本量需要的拨号次数

在第 2 章,我们讨论了为一个特定的调查建立的抽样库究竟应该有多大,但这一数字不等于调查员为最终达到调查规定的样本数而需要拨叫的电话号码数。例如,格罗夫斯和卡汉(Groves & Kahn,1979)在他们很有创造性的美国本土 RDD(一阶)抽样调查中,最终完成的访谈即样本量是 1 700,但总共处理的电话号码(即呼叫单)有 13 000 个,而这些电话号码一共被拨了 44 000 次(即有 44 000 个处理结果)。

为了帮助读者更为深刻地了解需要调查员拨叫的电话数，我们来看一下 NUSL 于 1992 年处理的一个全美二阶 RDD 抽样库。这一调查最终完成的访谈数为 1 500 个，而整个调查联系过的住户总数和从抽样库中提取的电话号码的总数，分别高达 2 400 户和 4 600个。表 3.2 列出了在调查进行的 6 个星期中，总共约 18 000 次拨号得到的具体结果。可能与所有的 RDD 调查一样，在这一调查中出现频数最高的结果是通了但没有人接，它几乎占了所有拨号次数的 1/3(34%)。而其他经常出现的结果则包括：住户应答机(22%)；接通但不是选定的被调查人(14%)；完成调查(8%)；未运行号码(8%)和住户拒绝(这一调查对 5% 的拒访做了拒访转变尝试)。表 3.2 也显示，拨号导致诸如非住户、回答人拒绝、完成部分调查和语言障碍等这样的处理结果相对较少。

表 3.2　1992 年全美二阶 RDD 抽样拨号处理结果分布($N = 17\ 962$)

处理结果	占所有拨号数的百分比
未运行	8.0
非住户	3.8
从来没人接听(无应答机)	34.0
无成年居民	0.1
非英语住户/被调查人	0.7
住宅应答机	22.0
住户拒绝	4.6
接通，但未能与被调查人接触	14.2
入选的被调查人拒绝	1.9
完成部分调查	1.9
精神/生理残疾	0.4
完成调查	8.4

资料来源：西北大学调查实验室，1992

我们也可以用另一种方法从质量和数量意义上去理解抽样过程。具体的做法是：在 1 次、2 次和更多次的拨号之后，计算究竟有

多少张呼叫单完成了它的最终处理结果。表 3.3 便列出了于 1992 年举行的全美电话调查得到的结果。注意,在 4 579 张呼叫单中,只有 2 158(47%)张,拨了一次号就完成了它们的最终处理结果。如果我们将这些一次拨号就完成了调查的呼叫单从所有的呼叫单中去除,那么其余 2 421 张呼叫单的拨号平均数为 7 次。这一数字清楚地显示处理这一抽样库所需的总拨号量大大高于 10 年前进行类似的调查所需的总拨号量。造成这种差异的主要原因在于住户应答机的使用和其他比较难以联系的被调查人的存在。

表 3.3 所列的数据使我们可以从另一个角度(完成特定的调查所需的呼叫次数)进一步理解抽样过程。在全部 1 500 个最终完成的调查中,有 1/4 左右(27%)在第一次呼叫时就完成了,而另外的约 1/5(21%)则是在第二次呼叫时完成的。因此有一半左右的完成的调查只呼叫了一或两次就完成了。假如这一调查在试呼三次之后便停止了对呼叫单作进一步处理(就如许多市场调查所做的那样),几乎有 40% 后来逐渐可以完成的调查就将丢失。如果调查对任何电话号码都不作 5 次以上的呼叫,那么相当完成总数 1/4 左右的完成数将丢失。而对呼叫单作 10 次以上的处理,则可使丢失数下降到不足完成总数的 10%。

用表 3.3 所列的信息,我们可以计算出,在 18 000 次左右的呼叫中,约有 1/6 为第 10 次及以上的呼叫。如果回呼的目的是为了降低总调查误差,那么我们不禁要问:花费如此多的人力物力(即再作 2 863 次呼叫次数在 9 次以上的额外的呼叫)才完成 132 个调查(即总完成数的 9%)是否值得?这就是说,我们花费了很多的资源才使样本量有少量的增加(因而使抽样误差略有下降)。如果我们改而考虑降低总调查误差中的其他误差而不是抽样误差,那么同样的资源产生的效果是否会更大呢?

表 3.3 所有呼叫单（$N=4\ 579$）中，所有最终处理结果为完成调查（1 500）的呼叫次数分布

呼叫次数	所有呼叫单			完成调查		
	频数	百分比	累计百分比	频数	百分比	累计百分比
1	2 158	47.1	47.1	408	27.2	27.2
2	609	13.3	60.4	307	20.5	47.7
3	386	8.4	68.9	196	13.1	60.8
4	240	5.2	74.1	180	8.6	69.4
5	210	4.6	78.7	117	7.8	77.2
6	150	3.3	82.0	69	4.6	81.8
7	114	2.5	84.5	59	4.0	85.7
8	90	2.0	86.4	42	2.8	88.5
9	74	1.6	88.0	39	2.6	91.2
10 ~ 14	304	6.7	94.7	84	5.6	96.8
15 ~ 19	113	2.5	97.2	37	2.5	99.3
20 ~ 24	104	2.3	99.5	6	0.4	99.7
25 或以上	25	0.5	100.0	5	0.3	100.0
总计	4 579	100.0		1 500	100.0	

资料来源：西北大学调查实验室，1992

RDD 样本中最终处理结果分布

表3.4展示了用于1992年RDD调查的每张呼叫单的最终处理结果——取自抽样库的每张呼叫单在实地调查结束时的实际状态。在进行如表所示的17 962次呼叫过程中,共使用电话号码4 579个。就二阶RDD抽样方法的效率而言,一半以上的电话号码接通了住户,其中频数最高的最终处理结果是完成调查(33%),而其余发生频数较高的最终处理结果码是未运行号码(29%)、非住户号码(15%)及拒访与完成部分调查之和(13%)。从无法联系因此有潜在的无回答误差的角度看,在所有处理过的电话号码中,无论以被调查人从未在现场(2%)——其中包括某些住户应答机,还是以从未有人接听(4%)结束的号码所占比例都比较小,而且其中大多数很可能根本不是居民住户(参见 Groves & Lyberg,1988,p. 199)。

表3.4 1992年全美二阶RDD抽样调查拨号处理结果分布(N=4 579)

处理结果	占所有拨号数的百分比
未运行	29.1
非住户	14.8
从来没人接听(无应答机)	4.0
无成年居民	0.5
非英语住户/被调查人	2.3
住宅应答机,从来无人接听	1.0
住户拒绝	6.6
未能与被调查人接触	2.1
入选的被调查人拒绝	4.6
完成部分调查	1.5
精神/生理残疾	1.0
完成调查	32.8

资料来源:西北大学调查实验室,1991

另一个 RDD 实例

我们的某些发现来自一个于1991年在芝加哥地区进行的规模比较大的RDD调查。在该次调查中,我们采用了以每一前置号和去除了每一前置号中的空后置号分层的一阶RDD抽样(参见第

2 章)。在 1992 年进行的全美电话调查中使用的处理结果类别就是该调查使用的类别的改进型。1991 年的调查也是由 NUSL 组织实施的,共处理呼叫单 3 005 张,最终完成 1 027 张。

表 3.5 列出了 1991 年芝加哥地区调查使用的所有的呼叫单的最终处理结果,与 1992 年的全美调查(见表 3.4)相似。完成调查的号码占所有来自抽样库的号码的 1/3 左右(34%)。而未运行(21%)、非住户(16%)、拒访和完成部分调查(17%)则是在完成调查之后,其他几类出现频数较高的最终处理结果。同样,无论以被调查人无法联系到(3%),还是以从来没有人接听(3%)告终的电话号码,在所有处理过的电话号码中所占比例也都比较小。

表 3.5　1991 年芝加哥地区分层 RDD 抽样,呼叫单($N=3\ 005$)最终处理结果

处理结果	号码数	占所有样本号码百分比
未运行	636	21.2
非住户	478	15.9
从来没人接听(无应答机)	78	2.6
无成年居民	78	2.6
非英语住户/被调查人	98	3.3
住宅应答机,从来无人接听	15	0.5
住户拒绝	302	10.0
未能与被调查人接触	69	2.3
入选的被调查人拒绝	150	5.0
完成部分调查	28	1.9
精神/生理残疾	39	1.3
完成调查	1 027	34.2
其他	7	0.2

资料来源:西北大学调查实验室,1991

总之,RDD 抽样调查中监控者和调查员的工作量将取决于某一特定的调查所涵盖的地理区域和抽样库的生成方式。如果在建立抽样库的过程中,我们采用了有关抽样区域内每一前置号的十分充分的信息或二阶抽样法,那么便有希望使调查过程中所需处理的电话号码数比简单(未分层)RDD 抽样减少 50% 以上。

以清单为依据建立的抽样库

如果抽样库是以一份清单或电话号码簿为依据建立的,那么

为完成规定的调查数所需处理的电话号码，较之用 RDD 建立的抽样库少得多。然而，即使使用最新的清单，实际处理的呼叫单数比计划完成的总数多 50% ~100% 也是不足为怪的。因此，一个计划完成 400 个访谈的调查，往往可能需要处理一个有 600 个电话号码或更多的抽样库。

下面的例子是我所经历的一个典型的用清单为依据建立的抽样框（库）对特定总体，即对执法人员（警官和州治安官员）进行调查的例子。它于 1988 年由 NUSL 实施。该抽样库是采用随机方法，从一个由普查局建立的全美所有执法机构的数据库生成的（参见 Lavrakas & Rosenbaum, 1989）。如表 3.6 所示，在完成对 788 名被调查者的调查过程中，共从抽样库中提取了 1 117 个执法机构的电话号码。（在实际拨号呼叫之前，事先给所有 1 117 个执法机构的主管都发了信，参见本书第 4 章。）与对公众进行调查的 RDD 调查相比较，该调查的最终处理结果中完成调查的比例是公众调查的两倍（71%）。与 RDD 调查不同，在各种情况导致的无回答中，无法联系的比例（16%）高于拒访（6%）。（尽管有人认为在组织实施非常完善的特殊群体的电话调查中，如果调查内容与被调查者自身的专业有关，而且致电进行访谈的时间又是被调查者比较方便的时间，那么被调查人一般都比较乐于参与调查。）

表 3.6 呼叫单最终处理结果分布（N = 1 117）1988 年全美执法机构主管抽样调查

处理结果	号码数	占所有样本号码百分比
电话号码不正确	30	2.7
无执法职能	39	3.5
无法联系	7	0.6
无已任命的首席执行官	9	0.8
部门拒访	22	2.0
首席执行官始终没有时间	174	15.5
首席执行官拒访	41	3.7
完成部分调查	3	0.2
完成调查	788	70.5
其他	4	0.4

资料来源：Lavrakas & Rosenbaum, 1989

同组追踪调查(Panel Studies)

调查研究人员和抽样库监控人员可以进行同组追踪调查,即在相隔一定时间之后重访同一组调查对象。这样的调查中应抱什么样的希望呢?我自己的经验显示,调查员是否能顺利地重访原来的被调查者(特别是在 RDD 调查中),很大程度上取决于抽样区域内居民的迁移模式(参见 Lavrakas, Settersten, & Maier, 1991)。在大城市中,如果我们在时隔一年之后进行重访,大约可以找到 50% ~60% 的原来的被调查者,而在农村地区这一数字比较高,约为 70% ~75%。

注意,通过电话进行重访,并不一定需要知道原被调查者的姓名。就成功的可能性而言,知道原被调查者的姓名与知道原被调查者的性别和年龄但不知道他们的姓名相差无几(参见 Lavrakas 等,1991)。在同组追踪调查的第二轮中,仍然可以联系到的原被调查者,能够完成调查的比例高达 90%,其余 10% 则可能拒访,或在调查期间不在现场。

回呼和无回答

依据计划或预算规定的回呼次数的不同,一般公众和特定总体的电话调查的最终处理结果的分布方式将有所不同。许多调查研究者认为,在市场研究中,三次回呼便已足够精确了。相反,在学术性调查和大多数联邦政府资助的调查中,一般则需要回呼 10 或 20 次甚至更多次(参见 Traugott, 1987)。不言而喻,后者的最终处理结果分布将明显不同于前者。在我们自己做的有关 RDD 调查中回呼问题的多变量分析显示:那些只需少数试呼便可以联系上的人,通常为女性、文化程度较低、没有工作、已婚和年纪较大的人。此外体弱多病的人也比较容易联系上(Merkle, Bauman, & Lavrakas, 1991)。这一结论与其他有关易于联系和难于联系的被调查人研究得到的结论是相当一致的(参见 Groves & Lyberg, 1988)。如果诸如这样的因素在一定程度上与调查实际关注的问题有关,并且在调查执行过程中回呼的次数又相对较少,那么在调查的最终结果中几乎可以肯定会有无回答误差(偏倚)存在,这样

的偏倚无论大小,一般都不十分明显。正因为如此,每一个调查研究者都必须事先考虑到多次呼叫的相对重要性,并就资源的合理配置做出正确的决策。

抽样库处理结果及调查的回答率和无回答率

电话调查中之所以要采用可高度控制的系统处理抽样库的主要原因是,它能使回答率在给定的抽样区域内达到相当高的程度,从而减少无回答误差发生的可能。较之类似的 PAPI 式调查,在使用软件来控制抽样库的 CATI 调查中,为了进一步降低无回答率,调查研究人员需要花费更多的精力和时间关注整个处理过程,在遇到难以处理的被调查人或住户时,情况尤其如此。

我们并不准备在这一节全面地阐述电话调查的回答率和无回答率问题,只是对其中的某些内容作一些扼要的介绍。无论是准备撰写一个关于已经完成的调查所采用的调查方法的报告,还是评估另一个调查有可能达到的质量(效度),这些问题都是必须加以考虑的。近来,专门从事调查研究的专业人员都认为,最好把回答率视为一种值域而非一个单独的数值。一般来讲,影响回答率的因素有调查的题目、问卷的长短、操作实施调查的组织者和调查员的才能、实地调查时间的长短、回呼和转变拒访的实施规则等。此外,格罗夫斯(Groves,1989,p. 133)提醒我们,只要那些接受了调查的人和没有接受调查的人在调查试图测量的变量上可能存在着差别,那么这些比率本身就不是无回答误差的直接量度,它只是回答或无回答率的函数。希望对这些问题有进一步了解的读者,可参阅福勒(Fowler,1993)、弗瑞(Frey,1989)、格罗夫斯和莱伯格(Groves & Lyberg,1988)的有关著作。

为了能在这一章对这一问题做较为深入的讨论,我们将进一步使用前面已经使用过的两个例子(1991 年在芝加哥都市地区进行的 RDD 调查和在 1992 年进行的全美 RDD 调查)。表 3.4 和表 3.5 分别列出了这两个调查使用的呼叫单(即电话号码数)的具体分布。

抽样库效率

表3.7列出了与这些调查有关的各种回答率。回答率的各种计算方法取决于用什么作分子和分母(参见 Groves,1989;Kviz,1977;Traugott,Groves,& Lepkowski,1987)。例如,在1992年全美RDD调查中,量度抽样库接通运行号码的效率值为71%(即3 246/4 579=0.709),因此只有1 333个或29%的取自抽样库的电话号码被确认为未运行号码。就这一点而言,这一抽样库是一个效率很高的抽样库,它也再一次证实了米托夫斯基-瓦克斯伯格二阶抽样可以提高抽样库效率的观点是正确的。而1991年的地方性调查的调查效率则更高,达到了79%,也再一次证明在地方性的RDD调查中,用前置号和去除了空号后的后置号进行分层是行之有效的。

作为1992年抽样库接通居民住户效率的量度,若采用保守的估计方法,应从分子中减去676张接通非居民住户的电话号码的清单,但仍应保留184张先后呼叫过30次以上而始终无人接听的电话号码的呼叫单。因此这一效率的估计值应为2 670/4 579即56%。不过经验告诉我们,在所有始终没有人接听的电话号码中,十有八九都是非居民住户的电话。这一调查的实地调查进行了7个星期。在这一期间,所有这些号码,在每星期的不同的日子或周末的白天或晚上的不同时间总共呼叫了30次或更多次。因此我们有理由认为,这些电话号码大多数都是大型超市或其他非居民居住区的投币电话,或者当时根本就没有运行的电话。基于这一推理,我们可以将这184张呼叫单从分子中去除,从而得到抽样库接通这一抽样区居民住户的效率的更为保守的估计值52%。1991年调查的类似的估计量之值为60%。

完成率

为了计算1992年调查的粗完成率,我们可以求1 500张完成调查的呼叫单与整个调查使用的全部4 579张呼叫单之比,得到的比率约为33%。而在1991年的调查中,粗完成率的估计值几乎与之相同(34%;1 027/3 005)。

表 3.7　1991 年和 1992 年调查完成率计算方法及完成率一览表

比率类型	计算方法	1991 年地方性调查		1992 年全国性调查	
		分子/分母	比率	分子/分母	比率
抽样库粗效率	运行号码/全部号码	2 369/3 005	78.8%	3 246/4 579	70.9%
抽样库接通住户效率	可能的住户号码/全部号码	1 891/3 005	62.9%	2 570/4 579	56.1%
抽样库接通住户效率的保守估计值	非常可能的住户号码/全部号码	1 813/3 005	60.3%	2 386/4 579	52.1%
粗完成率	所有完成的号码/全部号码	1 027/3 005	34.2%	1 500/4 579	32.8%
保守完成率	所有完成的号码/可能的住户号码	1 027/1 891	54.3%	1 500/2 570	58.4%
最合理的完成率	所有完成的号码/所有合格的号码	1 027/1 598	64.3%	1 500/2 216	67.7%

资料来源:西北大学调查实验室,1991,1992

1992 年调查住户电话号码访谈更为有效的成功率的估计值，应使用所有可能的住户号码(包括始终无人接听的号码)作分母。这一比率为 1 500/2 570 即 58%。1991 年调查的这一比率略低(54%;1 027/1 891)。更为合理的成功率估计值，应从分母中去除那些已经证明是不符合调查要求的，即不合格的电话号码。这样的电话号码包括:户内不存在任何一个能满足调查规定的年龄要求的被调查人;指定的被调查人不会说英语，或因永久性残疾无法参与调查和在许多次试呼之后，始终无人接听等这样的住户电话号码。我们认为这样计算得到调查的成功率最为合理，在 1992 年调查中，这一数字为每 3 个合格的被调查人中就有 2 个完成了调查(68%;1 500/2 216)。而在 1991 年调查中，这一数字则略低(64%;1 027/1 598)。

而总的来讲，我们究竟可以在这两个调查的抽样库的比较中了解到什么呢？我们认为，尽管两个调查各自使用的方法，就满足各自特定的调查目的而言，也许都是最好的，但从接通合格的住户这一角度看，1991 年地方性调查使用的方法比 1992 年全国性调查的效率似乎更高一些。而从住户的访谈成功率，也就是从降低无回答率的角度看，1992 年调查的效率则稍高一些。(注意，1992 年聘用的调查员的能力稍高于 1991 年，这有可能使 1992 年调查的无回答率保持在较低的水平上。)最后，1991 年和 1992 年调查的拒访转变率分别为 27% 和 25%。

无回答率

其他经常计算的各种比率都与无回答的性质有关。表 3.8 列出了 1991 年和 1992 年调查的接通(或可能接通的)但最终未能完成调查的住户电话号码的各种中止处理结果码的详细情况。迄今为止的所有调查，构成无回答问题的最主要因素是拒绝回答，它也是在公众电话调查中经常出现的问题(Groves & Lyberg,1988)。在两个调查中，大多数拒访都来自住户中某一个并非指定的被调查者的人，或先于选定的被调查人接听电话的人。在两个调查中，大约 1/10 的拒访是因为选中的被调查人存在语言障碍或在呼叫时始终不在调查现场。在 1992 年的调查中，因始终无人接听而导致

的无回答的比例是1991年的两倍。最后,我们认为还有一个值得读者注意的问题是,两个调查都只有非常小的一部分无回答是由应答机所致。

表3.8 1991年调查(N=779)和1992年调查(N=1 049)无回答[*]原因构成

无回答原因	未完成百分比	
	1991年调查	1992年调查
始终无人接听(无应答机)	10.0	17.5
非英语住户/被调查人	12.6	9.9
始终未与选出的被调查人取得联系	8.9	9.1
选定的被调查人拒访	19.2	19.7
访谈中断	3.6	6.7
住户应答机/始终无人接听	1.9	4.2
被调查人精神/生理残疾	5.0	4.3
住户被调查人拒访	38.8	28.6

数据来源:西北大学调查实验室,1991,1992

* 这些都是已经确定未完成的访谈,最后一次呼叫的处理结果。

这两个调查中的拒访/完成比率,可被称之为不合作率(参见Groves & Lyberg,1988,p.200)。它可以从计算完成调查的电话号码的呼叫单数与以某种形式的拒访(即住户拒访、被调查人拒访、完成部分调查)结束的电话号码的呼叫单数之比求得。在1991年的调查中我们可以用1 507(1 027+480)作分母而用480(包括所有形式在内的拒访数)作分子,求得这一比率约为32%。1992年调查的这一比率约为28%。

大家应该记住,样本的无回答率只不过是调查中的各种形式的无回答中的一种。在这一节讨论的各种无回答问题中,除了样本无回答问题之外,还有问卷各个题项的无回答问题(即丢失某一特定题项的答案)。令我们感到庆幸的是,如果我们能在电话调查中聘用比较能干的调查员,并使用非常完备的调查工具,那么对绝大多数题项而言,题项无回答不可能对总调查误差产生任何明显的影响。

用加权修正无回答误差

后抽样的统计修正(加权)是一种普遍使用,但并非总是精确

和有效的，试图对可能的无回答误差作一些弥补的统计方法（参见 Henry，1990）。一般地，这种方法主要用于样本的人口学特征值做后抽样的修正。修正通常都参照目标总体的同一人口学特征值（这些特征值一般都来自人口普查的数据）进行。例如，在某一次调查的样本性别构成中，男性所占的比例可能仅为 40%，而在总体中成年男性的这一比例却为 46%。这样每一男性被调查人将得到性别权数 46/40 或 1.15（即每一个男性将被当作 1.15 个被调查人，而每一女性则被当作 54/60 即 0.9 个被调查人）。后抽样修正通常基于一个未经证明的假定：那些“未加权”（注：原文为 upweighted，似是 unweighted 之误）的对丢失的被调查人作补救的人所提供的数据，与原来打算从其他与他们具有相似的人口学特征，但是“丢失”的人那里搜集的数据相似。那些对这一问题感兴趣，并且希望对它有更深入的了解的读者，可参阅格罗夫斯的著作（Groves，1989，pp. 156～188）。

注意，为了对无回答作补救而进行的统计加权，不同于规范和恰当地进行的对选择的不等概性（例如在 RDD 调查中，某些住户因有多个电话号码而具更大的入选样本的概率）作修正的后抽样加权。

小 结

在实地调查结束的时候，我们可以用各种各样的方法对抽样的质量进行检验。为了了解样本对总体的代表性，我们可以计算反映调查成功程度的各种相对比率，其中比较好的方法是计算调查的回答率和无回答率。那些准备出版发表他们调查结果的人，应当在他们的调查报告中报告这些比率，并清楚地说明范围误差和无回答误差对调查结果可能具有的影响，以便他人对已经完成的抽样的相对质量做出正确的判断。

混合式调查的抽样库

混合式调查应用的增长（参见 Dillman & Tarnai，1988；Traugott

等,1987)在一定程度上遏制了许多实际调查的回答率的下降。目前调查研究者也许正处在这样一种情况:他们在做某项调查时,部分抽样库以电话处理,部分抽样库以邮寄方式处理,而另一部分抽样库则使用个别访谈处理。这样的调查要求掌握用邮寄和个别访谈法处理的那部分抽样库的居住地址。这些地址也许包含在一个初级抽样框内,或可能需要事先通过电话调查搜集。当然,假如我们的调查问卷是用电传来"邮寄"的,那么就不必事先知道被调查人的地址。

注意,在过去的10余年间,存在一种不是很强却在持续增长的趋势,那就是选定的被调查人拒绝通过电话进行调查,但乐于填写邮寄给他们的问卷。根据我自己的经验,这样的情况大多发生在老年人身上。因为许多老年人都曾经接到过来自老年杂志要警惕电话调查的警告。遗憾的是,要满足这样的要求并非像得到一个地址将问卷寄出那么简单。例如借助电话,特别是CATI式电话调查的问卷格式很少适合由被调查人自行填写。因此,一个单一模式的调查一旦开始之后,就很难立即转换成混合模式的调查。而那些事先就计划要做混合式的调查则有一个特殊的优点,那就是它使调查者可以事先寄出安民告示式的信件(参见第4章),提醒被调查人在不久的将来要对他进行一次调查,进而为未来的调查建立起某种信任。在业务性调查中,如果事先已经知道,或通过事先的电话接触知道了被调查者的传真号码,这样就可以在调查开始之前的适当时间用传真发去一份安民告示式的信件。

下面给读者介绍一个NUSL用邮寄、电话和个别访谈三者处理的抽样库实例,以帮助大家对在公众调查中可能遇到的混合式调查有进一步的了解。该调查于1991年进行,调查目的在于精确地测算地方校区(Morton Grove, Illinois)将来的招生人数。抽样设计要求对该校区内3 000个住户中的2/3进行访谈,而且回答率必须达到90%。初级抽样框由该校区内的所有邮政地址清单构成。我们首先对初级抽样框进行了校验,发现清单上10%左右的地址实际位于贯穿整个社区的街道上。当然通过这一校验,我们也确认了更多的地址。然后我们又用反向电话簿对抽样框作进一步的校正和补充,发现抽样框中约80%的住户在电话簿上登录了他们的

电话和地址。然后我们建立了一个含 2 210 个地址的系统随机抽样库，并给每一个地址发去了一张明信片，提醒抽样库中的所有住户，我们即将对他们进行一次调查。大约在一星期之后，给他们邮去了含一份 2 页长的问卷和贴上邮票的寄回填写后问卷的信封的邮件。10 天之后，所有那些尚未将问卷寄回且在清单上列出了他们的电话号码的住户，都会接到追踪电话。这个电话提醒他们协助填写问卷，或请求对方在电话上完成调查。而那些既没有寄回问卷，也没有在清单上登录他们的电话号码的住户，则会收到我们寄去的第二份问卷。随后进行的便是第二轮追踪电话。最后，我们对所有仍未完成调查的住户都进行了个别联系。

混合式调查抽样库处理历时 8 个星期，在实地调查最后完成时，回答率高达 92%，其中通过邮寄、电话和个别访谈完成的分别占 66%、15% 和 11%（就与抽样库中 8% 的未能顺利完成调查的住户相关的潜在的无回答误差而言，我们有理由认为无回答户中的大多数，都是老年住户和其他无学龄儿童的住户）。

在这样的混合式调查中，电话调查无论是作为邮寄调查的辅助手段，使邮寄调查能更顺利地完成，还是作为邮寄调查的替代，直接完成全部调查中六分之一的问卷调查都可以很好地发挥自己的长处。而从费用和效益比的角度看，尽管混合调查的时间是单一的电话调查的 2 倍，实际费用却几乎没有什么差别。而混合式调查非常有可能显著减少只用电话调查而产生的未覆盖误差和无回答误差的优点，足以抵消它调查时间较长的不足。

遗憾的是，有利于这一混合式校区研究取得成功的条件——特别是与被调查者密切相关的非常短的问卷并非常常存在。然而因为混合式调查可能具有的优点，所以调查研究者在进行研究设计时，应该考虑周全，不要错过每一个可以采用混合式调查的机会。

用 CATI 系统处理抽样库

正如我们前面所说，并非所有 CATI 调查都采用计算机来控制

抽样库,这也就是说,有些CATI调查,只用计算机来管理问卷,而抽样库的处理仍然依靠人工,采用纸笔及呼叫单等方法。

尽管CATI调查有诸多引人注目的优点,但就抽样而言,其最吸引人的地方是将在PAPI调查中需要由监控人员人工完成的抽样库处理工作,改由计算机完成。这就使监控人员能有更多的时间投入到抽样库处理和调查员督导上,对遇到的各种问题做及时和正确的处理(参见第6章)。这样,未覆盖误差、无回答误差和测量误差就有可能降低,从而使总调查误差也减少到更低的水平。

练 习

1)用表3.1所列的处理结果作例子,设计一种新的分类系统,该系统应使调查员使用的处理结果码为10类,或少于10类。

2)编写一篇标准化的留言,以作调查员在应答机上的留言之用。字数不得超过50字。

3)编写两段电话调查员可能经历的对话,说明住户拒访和被调查人拒访之间存在的差别。

4)用你们的教授给你的数据计算表3.7所列的所有比率。

5)用你们的教授提供的数据,编制一张如表3.8所示的未完成调查登录表。



选择被调查人并确保其合作

Selecting Respondents and Securing Cooperation

有些对电话调查不太了解的人错误地认为,第一个接听调查员电话的人就是应该进行调查的被调查人。在任何设计用于搜集目的总体中有代表性的单位内样本(如住户内的成年人),从而降低未覆盖误差发生的可能性的调查中情况并非如此。例如,虽然婴儿的性别比例几乎为50:50,但大多数城市的成年人口的男女比例为45:55(男性:女性)。一个性别构成的有代表性的抽样调查,必须采用一种系统的被调查人选择法,以保证样本中被调查人性别比例的平衡。

本章讨论的各种问题,都与如何在一个抽样单位(如在一个住户内)选择有代表性的被调查人,以及如何保证入选的被调查人的合作(从而减少无回答误差)有关。在与住户第一次接触时向他们介绍调查的方式是调查能否顺利进行,特别是顺利地完成被调查人的选择的关键。正因为如此,在本章我们也会对编写简明的调查介绍/被调查人选择顺序的重要性予以特别的关注。

本章首先介绍有控制(系统)地选择电话调查的被调查人的重要性。此后我们将讨论在接触之初,调查员如何通过只言片语的介绍,迅速地与对方建立友好关系的重要性。这一部分内容包括:为什么必须对调查员进行培训,以使他们学会如何用标准化的答案回答被调查人经常问及的一些问题。在完成这一部分内容的介绍之后,我们将会扼要地回顾一下几种经常使用的选择被调查人的方法。最后,在即将结束本章的全部内容之前,我们将会对如何与被调查人在实地调查开始之前进行联系,以及这样的联系与被调查人的选择和保证取得他们的合作的关系等问题展开一些讨论。

争取良好的开局

在微芯片时代,不断发展的高级电话设备遍及整个社会,所以许多人常常抱怨受到来自素不相识的人的电话的打扰,也就不足为怪了。在今天,电话推销已经成为各种买卖销售自己产品,或搜集各种信息的成本效益显著的方法。有些电话推销,已经不需任何人工,完全通过机器进行。而这一切就是严肃的电话调查所无

法摆脱的"真实世界"的社会环境。在这样一个环境中,电话调查的调查员在电话接通以后,必须立即面对的挑战是如何给对方传达一个明确的信息,使对方明白我们的调查与他们有时接到的"垃圾电话"是不同的。而这正是我们编写调查简介/被调查人选择顺序等引导性用语最主要的目的。调查员的责任就是将它传达给对方,以达到取得对方合作的目的。

与个别访谈式调查不同,电话调查的拒访往往发生在联系刚刚开始,而非被调查人已经选定,或已经开始一一宣读问卷中的问题的时候(参见 Collins 等,1988)。奥森伯格和卡乃尔(Oksenberg 和 Cannell,1988)指出,大约 40% 的拒访,发生在调查员刚开始作简介时,另有 50% 发生在简介/选择顺序进行过程中,而还有 10% 或更少的拒访,发生在调查员已经开始问卷调查之后。因此电话接通以后的最初 30~60 秒,特别是从尽可能地降低无回答误差的角度看,是电话调查能否成功的关键时段。

那么调查研究人员应该如何使在调查员和被调查者接触之初都有可能发生的两类误差(未覆盖误差和无回答误差)同时达到最小的程度,即使二者保持平衡?一般来说,调查研究者都希望在实地调查结束时,得到的最终样本能在最大程度上具有人口学上的代表性,以避免未覆盖误差的可能影响。然而,那些可能用于选择抽样单位内个体的最有人口学代表性的混合式方法,通常也都是一些很费时间的方法。它们往往会对被调查人有较大的干扰,令他们感到厌烦或疑虑,进而拒绝接受调查,导致无回答误差的上升。

调查员调查技术的高低,固然是在这两个彼此冲突的目标之间保持平衡,从而使电话调查圆满成功的决定因素。然而,即使是最优秀的调查员,在面对一份措辞不当、用语含混不清甚至粗俗的简介/选择顺序时也会无计可施。相反,只要训练得当,加之一份用语亲切动人的调查简介和逻辑清楚的被调查人选择法,即使经验不多的调查员也能使调查圆满成功。

显然,在抽样以清单为依据,而且被调查人的姓名也已经清楚的情况下,调查员在选择被调查人时所需要做的事,只是找到那个已经确定了的被调查人而已。但是在许多用清单和所有的 RDD 抽样中,我们并不知道抽样单位(例如住户)中,那个应该被我们调

查的人的姓名。(有一个例外,那就是通过以前没有取得成功的试呼,我们已经知道了指定的被调查人的姓名。)因此,一个设计用于搜集个人层次的总体参数(相对于住户层次的量度)估计值的调查,必须采用一种系统的选择方法,以尽可能地提高调查数据的外部有效性(即可概括性)。

诚如前述,那些只调查第一个接电话,或听起来年龄似乎已经达到可以回答问卷上的问题的人的调查是不可能搜集到能有效反映目标总体的态度、行为和经验的数据的,因为这样做有可能在抽样单位内产生未覆盖误差。那些允许调查员愿意在抽样单位内调查什么人就调查什么人的调查,也会有同样的问题。可以设想,如果我们允许调查员只是调查抽样单位中愿意和他交谈的人,那么许多调查员都愿意这样来选择他们的被调查人。而用这种方法选择被调查人很可能导致最终的样本没有什么代表性。因此,被调查人的选择必须采用系统且无偏倚的方法。这就是说,调查员不能自行其是地从抽样单位中任意选择被调查人。

抽样单位中的被调查人可以用真概抽样法(true probability sampling scheme)选取,尽管调查研究人员并非总是必须或希望采用诸如这样非常严格的抽样法,因为它过于费事,因而会给被调查方过多的打扰。而就大多数调查的调查目的而言,它们可以接受的是一种可平衡性别和年龄的系统选择法。因为大多数抽样单位(如住户)在许多其他的人口学特征上(如种族、教育和宗教)具有一定的同质性,有代表性的单位内抽样应该精确地涵盖总体中这些其他的人口学特征。

简单扼要地介绍一个调查

在有关调查简介应该含有多少信息量这一问题上,专业的调查工作者至今仍然是众说纷纭,莫衷一是。有关的文献也不能给我们提供这一问题的确切答案(参见 Groves & Lyberg, 1988, pp. 202~210)。而我个人则比较倾向这样的观点,顾名思义,简介应该简单明了,能有助于被调查人积极配合我们的调查,开始回答问

卷上的问题。但有时也会有例外,有些调查的引导语中必须对问卷的组织结构和类型特殊的问题做介绍。此外,对调查能否成功来讲,开场白的内容固然重要,但调查员是否能很好地运用它也许更加重要(参见第5章)。

我认为一篇简介应该有足够多的信息,以尽可能减少接听电话的那一方,在听到一个陌生人致电希望进行一个调查时,可能产生的任何紧张心理(参见 Frey,1989,pp. 125 ~ 137)。也就是说,尽快建立对方对调查员的信任是调查简介的主要任务。

与此同时,经验告诉我们,在访谈进入问卷调查之后,再得到被调查者的合作,甚至是充分的合作是比较容易的。这就好比一只脚已经进门,再迈进另一只脚就不难了。调查介绍越长,和那些有可能成为我们的被调查人的人必须被动地聆听的内容越多,他们在调查真正开始前便已经对调查失去兴趣的可能就越大(Dillman,Gallegos,& Frey,1976)。只要可能,尤其是在调查研究人员本身对调查简介的文字仍然心存疑虑时,调查简介应当与调查的其余内容,即被调查人选择方法和问卷内容本身同时进行试调查,然后再根据试调查得到的信息,对它做进一步的改进。

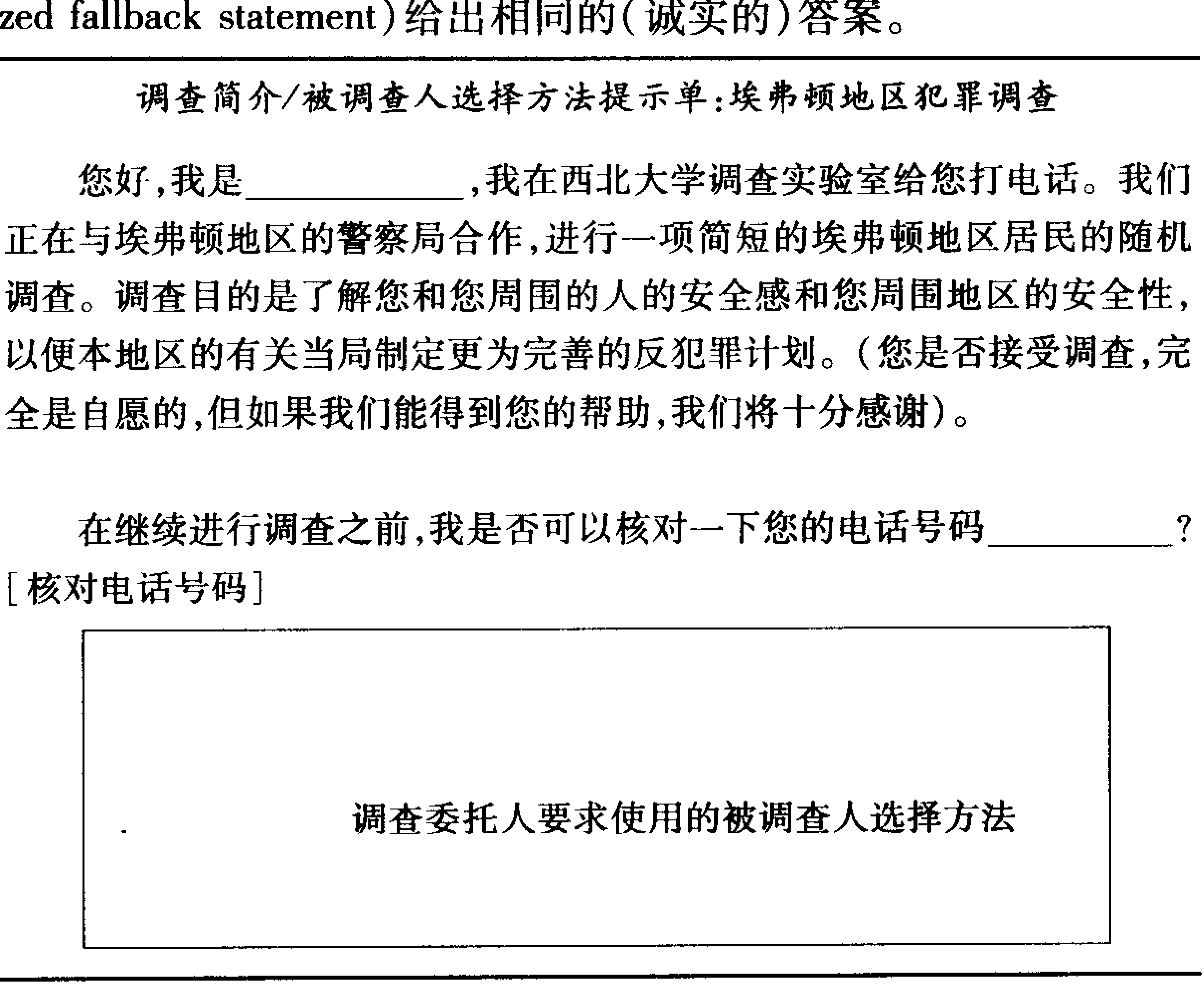
怎样编写调查简介

尽管在编写调查简介时,每一个调查研究者的措辞习惯可能各不相同,但以下内容则是美国公众舆论研究会(AAPOR,1991)和民意调查理事会(Gawiser & Witt,1992)一致推崇的。而我们则坚信,这些内容必须写入调查简介。

1. 调查员自我介绍、本人所属机构和调查资助机构;
2. 调查目的和抽样区(或目标总体)简介;
3. 一段诚恳和殷切的致词,吁请合作;
4. 核对调查员拨打的电话号码。

图4.1是NUSL使用的一页印有一般介绍用语的调查简介/选择方法的实例。在有些调查中,这样的简介必须附在每一张取自抽样库的呼叫单上,而在另一些调查中,则只需在调查站中放上一张诸如这样的调查简介就可以了。图4.1所示的介绍性陈述始于调查员的自我介绍,然后阐明了打电话的地点和打电话的原因。

其中有关打电话的目的介绍的那段文字,同时也对抽样区域和调查组织者的有关情况做了介绍。只要可能,我们也应该在这里含蓄或直接地说明调查资料的使用原则,除非这样做可能会使被调查人对调查产生疑虑。如果问卷比较短(例如只用 10 分钟),我认为还可在简介中简单提及调查需要的时间。如果访谈的时间在 15 分钟以上,我认为最好不要在简介中提及调查需要用的时间,除非被调查人问我们"调查需要多少时间"。在这样的场合,我们必须事先对调查员进行培训,使他们能用标准化的回撤陈述(standardized fallback statement)给出相同的(诚实的)答案。



调查简介/被调查人选择方法提示单:埃弗顿地区犯罪调查

您好,我是＿＿＿＿＿＿,我在西北大学调查实验室给您打电话。我们正在与埃弗顿地区的警察局合作,进行一项简短的埃弗顿地区居民的随机调查。调查目的是了解您和您周围的人的安全感和您周围地区的安全性,以便本地区的有关当局制定更为完善的反犯罪计划。(您是否接受调查,完全是自愿的,但如果我们能得到您的帮助,我们将十分感谢)。

在继续进行调查之前,我是否可以核对一下您的电话号码＿＿＿＿＿?
[核对电话号码]

调查委托人要求使用的被调查人选择方法

图 4.1 调查简介实例

调查简介中也应该包括对调查员拨叫的电话号码进行核对这一内容。这样不仅可使调查员与对方的联系进一步正式化,而且还有利于提高抽样库的处理和控制的质量。在本书的第 1 版中,我认为应在对方接听电话之后,立即核对电话号码(Lavrakas,1987,p. 84)。现在我认为,以这种方式开始调查未免有些唐突,许多被调查人也许难以接受。所以核对电话号码的工作可以稍晚一

点进行,在此之前,可以插入一些比较亲切的问候性的话语。核对RDD调查拨叫的电话号码的这一改动,无论就样本控制的目的,还是就避免电话转接所产生的问题而言都十分重要(例如电话转接会以某种我们不知道的方式改变所选择的概率)。

回撤陈述

多数调查也许并不需要做任何回撤陈述,因此我们不必多此一举,在简介中对调查的目的做过于详细的介绍。如果被调查人在开始决定是否参加调查之前,希望对调查有更多的了解,那么调查员就有必要根据事先起草的一份标准化的调查介绍,向被调查人宣读或解释。而对那些勉强参与调查的被调查人,我们则应通过培训使调查员有一定的临场判断力,使他们能在调查现场做出判断,如果确实需要,他们能给这些人就调查目的做出更为详尽的解释。

预期的被调查人主动想从调查员那里了解的问题,归纳起来不外乎几种基本的类型。但这样的情况只是偶尔而非经常发生的(参见 Lavrakas & Merkle,1991)。读者应注意,记住是偶然,而非经常发生这一点很重要。因为如果预期的被调查人经常向调查员问起某些问题,那么我们便应将这些问题的答案,事先编入简介中,然后由调查员向每一个被调查人宣读。有关这一问题的初步调查表明,有1/3左右的被调查人在调查开始之前,会给调查员提出许多某些类型的问题(Lavrakas & Merkle,1991)。(不仅如此,与白人调查员相比,黑人调查员被问到的问题更多。虽然差别不是很大,但其发展趋势却在统计上具有显著性。但除此之外,我们还没有发现其他的人口学特征,会对被调查人提出的问题有影响。)

然而在大多数情况下,调查员并不需要对被调查人做很详细的介绍。调查速度势必会因不必对每一个被调查人做详细的介绍而加快,故无回答误差发生的可能也会因此而减少。被调查人偶尔会问及的问题大概可概括为以下几种类型:

1. 调查的目的是什么和如何使用调查结果?
2. 你是怎样得到我的电话号码的?
3. 调查的实施/组织者是谁?

4. 为什么本户中的其他人不可以作为被调查人?

对上列每一类问题都必须编写回撤陈述,并提交给调查员,以使他们能对向他们提出问题的被调查人给出诚实的、标准化的回答。图4.2展示了一个典型的回撤陈述单的实例(读者也可参阅Frey,1989,p.230)。在所有情况下,对被调查人的解释都应包括这样的提示:被调查人给出的答案是会保密的,以及被调查人是否接受调查完全是自愿的,以使调查符合"知情自愿"的原则。

C25 调查员回撤陈述单

介绍调查:我们的调查非常简短,只需要二三分钟。大多数问题都与您对心脏移植手术所持的态度有关。我们调查的目的并不在于了解您个人是否了解心脏移植手术,而在于了解整个芝加哥地区的人们对于这样的手术的真实态度是什么。我正在西北大学修一门叫做"高级报告撰写技术"的课程,作为课程的作业,任课教授安排我们班的同学来进行这一调查。您是否愿意接受我们的调查,完全是自愿的,但是如果您能参加我们的调查,我们将十分感谢,因为它会对我们有很大的帮助。

调查资料的使用:作为我自己选修的课程的一部分,我必须给任课教授撰写一份调查报告。而我们整个班必须调查几百个芝加哥地区的居民。所有答案将被合并整理在一起,任何个人都可以对我们的问题不予作答。您是否愿意接受我们的调查,完全是自愿的,但是如果您能参加我们的调查,我们将十分感谢,因为它会对我们有很大的帮助。

随机数字拨号:您的电话号码是由随机数字拨号法选出的。我们的老师将所有代表在芝加哥地区运行的电话分局的3位数输入计算机,然后再由计算机随机生成7位电话号码中的后4位数。……这样生成的电话号码恰好包括了您的电话号码。……我们之所以采用这样的方法,是因为我们必须调查整个芝加哥地区的居民,不论其电话号码是否登录在电话簿上。而这样的方法,是我们惟一可采用的,能使调查结果可以充分代表芝加哥地区居民观点的方法。

性别选择:我手头上有一张单子,上面的内容告诉我应该对您家中的哪一个成员进行调查。有时它告诉我应该是家中年长或年轻的女性,有时它却告诉我应该是年长或年轻的男性。这样的方法是我们可以使年龄和性别保持充分平衡的惟一方法。

查询电话号码:如果您对本调查有任何疑问,您可以在白天致电491-5662,与我的老师拉弗拉卡斯博士联系。

图4.2 回撤陈述单示例

回撤陈述的目的在于帮助调查员向未来的被调查人说明，用一点时间接受调查是一件有意义（且没有任何坏处）的事。这一点，编写回撤陈述的人必须牢记在心。

有关被调查人的物质奖励问题

在草拟简介、回撤陈述和被调查人选择顺序的文字时，调查研究者都必须考虑是否给予（或不给予）参加调查的被调查人以一定的报酬。我们建议对这一问题感兴趣的读者可查阅格罗夫斯（Groves，1989，pp. 215 ~ 236）撰写的更为详细的有关社会科学的文献。例如有的调查研究者认为，可以考虑给被调查人付一些调查费，尽管在电话调查中，要做到这一点有一定困难。其他形式的物质奖励，如随研究报告送去一张彩票或其他什么赠品等也可考虑。有关物质奖励对调查的影响的文献可谓汗牛充栋，但有的观点相左、彼此矛盾，因此调查研究者使用时应加倍小心（Groyes，1989，pp. 215 ~ 217）。如果我们准备给参与调查的人一定的物质奖励，那么我们必须在调查简介中写明。

至于心理和社会奖励，调查研究者应该考虑在调查简介/选择顺序单中指出调查可能对被调查者带来的那种潜在的无形的好处。例如，表达自己观点的、从事一项感兴趣的工作（假定被调查者对调查的题目感兴趣）和有益社会（如果研究目的与社会政策有关）等等。在撰写回撤陈述时，研究者应该考虑人性中的其他方面，即利他和顺从。告诉有可能成为我们调查对象的那些人，他们正在帮助与他们联系的调查员，研究者正在期待他们的合作。回撤陈述的措辞中若能给对方提供在我们联系过的人中，绝大部分人都参加了调查（假定这消息对调查组织者是正确无误的）这样的信息，也会对调查的成功有所帮助。最后，我们也许应该在回撤陈述中加入一些文字，以帮助调查员应对那些有抵触情绪（比如说“还是让别的人来调查吧”这样的话）的被调查人。

CATI

控制引导性陈述的 CATI 系统给我们提供了对介绍性陈述进

行必要的剪裁,使之更贴切于调查对象个人的机会。迄今为止,我们还没有见到探讨个性化的陈述对调查的影响的文献。例如在同组追踪调查中,我们可以从先前的调查中得到被调查人个人的资料,因此我们在起草有关陈述时,将这些材料编入其中,使陈述更具个性化(参见 Lavrakas, Settersten, & Maier, 1991)。即使在 CATI 式的 RDD 横断面调查中,我们也可将那些在以前进行的未完成访谈的试呼中得到的调查对象的个人资料,编入用于以后试呼的介绍性文字中。

被调查人的选择和寻求他们的合作

即使那些控制不是很得力的电话调查,也会对在抽样单位内选择哪一位成员作为被调查对象这一问题有所考虑。(本节介绍的内容假定,调查员联系的那些人并非从一份清单或其他可以知道他们姓名的材料中选取的。)即使未经培训的调查员一般也都知道,可以作为调查对象的人必须在体力和智力上具有回答问题的能力。有些情况虽然不是十分普遍,但有时可能发生。譬如在一个讲英语的调查员试图和一个几乎听不懂英语的被调查人,或那些暂时或永远缺乏在智力上有效地回答各种问题的人进行调查时,势必会使调查的测量误差有所增加。本节所要讨论的内容,其目的并非想告诉调查员如何判断和防范这样一些误差。这些内容我们将在第 5 和第 6 章讨论。这一节所要讨论的问题是各种选择被调查人和如何寻求他们的合作的方法。

尽管我们从来都不希望读者在选择被调查人时,使用不加任何控制的方法。但我们认为从这种不加控制,任由调查员自行在抽样单位内选择被调查人可能产生的问题入手,开始进行被调查人选择问题的讨论却不失为一种明智之举。在下一节的讨论中,我们假定某一调查员正在进行一项要求被调查人报告他们自己对有关问题的态度、行为方式和经验的调查。如果调查的目的在于测量一些存在于抽样单位层次的某些设备的建构情况(如住户中

正在使用的电视数和其中使用有线电视的数量），那么户中的任何知情者都可以作为我们的调查对象，而无须控制未覆盖误差和测量误差。在这样的场合，被调查人的选择只是通过电话与抽样单位内的任何一个符合研究人员需要的人取得联系。（在这样的场合，一种最为简单易行且容易控制的方法是随机应变，这就是说在接听电话的是女人时，调查员就说“要调查户内的女性”，反之就说“希望调查户中的一位男性”。）

无控制选择

在采用无结构的户内被调查人选择方法的一般人口总体的调查中，户内任何一个合乎要求的人都可以成为我们的调查对象。在许多情况下，年龄是调查员选择被调查对象的惟一标准。例如许多调查规定，被调查人的年龄必须在 17 岁及以上。但一个有关男性对某些问题的观点调查的抽样则可能与之不同，因为调查员只须和户中当时在场的某一男性交谈，如果当时有男性在场的话。（在销售传统的“男性”产品的市场，如口嚼烟丝的电话市场调查中，经常采用无控制的户内成年男性选择法。）在其他一些调查中，也许要求调查员只对户主进行调查，而户主的定义则由被调查户自行认定。在这些调查中，正式的被调查对象的选择方法都没有打印在调查简介/被调查人选择单上，调查员只需在宣读调查简介之后，立即找户内任何一个合乎要求的人进行调查就可以了。

当调查目的是测量总体变量的存在水平时，无控制的被调查人选择可能产生两种类型的问题。第一种也是比较严重的问题是这种选择方法可能会在抽样单位内产生未覆盖误差，而未覆盖误差则有可能进一步导致调查的样本失去对目标总体的代表性。如果我们允许调查员从合乎要求的人中自行选择他们乐于选择的任何人，那么在由此而得到的样本中，那些在调查员打电话时更有可能在家的人的比例可能会大大高于实际的比例。在一般人口总体中，这些人通常是妇女和老人（参见 Salmon & Nichols，1983），如果打电话的时候有几个合乎条件的人在场，调查员就会自然而然地对其中一个正在现场，且（或）乐于接受调查的人进行调查。如果

调查中确实发生了这样的情况,那么在最后得到的样本中,年轻人和教育程度比较高的人的比例可能会偏高。遗憾的是,我们不能期待与可遇性(availability)和积极性(willingness)相关的偏倚能彼此抵消,从而使未覆盖误差自行消除。相反,由两种因素混合产生的未覆盖误差不仅是因调查而异和无法预测的,而且还不像在样本有高度的外部有效性时那样是可以避免的。

无控制选择产生的第二个问题可能会随时间的推移而增长。诚如我们所知,公众正在慢慢成为参与调查研究的老于世故的顾客,因而在对一般公众进行电话调查时,会有越来越多的人认为采用无控制的被调查人选择法选择被调查人是不够专业的。实际上,无论调查员还是调查研究人员,他们都不希望给潜在的被调查人留下参与他们的调查只是在做一件浪费时间而无任何实际意义可言的傻事这样的印象(即一个有非常大的总调查误差,因而可能没有任何效果的调查)。显而易见,一个有着不仅措词得当和没有任何强制性,而且显得很正式的被调查人选择顺序的调查一般比较容易为人们所接受。

有控制选择

在过去的30年间,在被调查人选择中采用的最为普遍的方法并非抽样单位内真概抽样法,而是一种不允许调查员自行其是地选择被调查人的系统方法。此外,这些方法也被设计成为在试图得到整个调查的被调查人在人口学特征平衡的样本的同时,最大限度地减少对被调查户(人)的打扰的方法。这些非概率的被调查人选择法,将在本节的后面部分讨论。现在,我们先来讨论标准的概率选择法。

凯思随机选择法

为大家公认的,在个别访谈中使用得最为严格的标准的被调查人选择法是凯思(Kish,1949,1965)发明的。这一方法也可用于需要得到在抽样单位内尽可能具有代表性的样本的电话调查。

即使采用凯思法仍然可能在抽样单位内存在某种较小的未覆

盖误差，了解这一点很重要。因为并非所有的抽样单位（如住户）都愿意或能提供全部合格者的清单。例如格罗夫斯和卡汉（Groves & Kahn，1979）在进行住户构成可信度检验时发现，约有9%的住户提供的住户成年人口信息与实际情况有差异，而这种差异大多因为户内还存在其他年纪较轻的成年人所致。除了有意误报导致的错误之外，并非所有的人都明白，某一特定的调查实际使用的选择合格被调查人的顺序的确切定义（参阅 Maklan & Waksberg，1988；Groves，1989，pp. 108 ~ 115）。这一问题在电话调查中出现的可能大于个别访谈，因为电话调查的调查员无法解决某些在个别调查访谈过程中，因能直接看见而可以立即解决的问题（Massey，1988）。

尽管凯思法在理论上被视为一种真概选择法，实际上在户内人口较多（>6 人）时，最年轻的被调查人的比例在样本中仍会略嫌不足。不过较之其他不严格的选择法，凯思法将最大程度地减少抽样单位内的未覆盖误差（尽管如下面将要指出的那样，它可能会使无回答误差有所增加）。如图 4.3 所示，调查员在宣读开场白之后，便应马上确认抽样单位内的全部符合条件的人数。在大多数情况下，这就意味着确定户内满足某种年龄要求的全部人数。而某些调查研究者希望调查员确认合乎条件的人与户主之间的关系，而另一些调查研究者则希望调查员了解所有合格者的姓名。无论采用哪一种方法，一般调查员都必须首先确定户主，然后再逐一登录所有的合格者。例如，假定一个调查需要抽取的调查对象是成年人，那么调查员一般可能首先登录丈夫和妻子，因为这两种人在美国住户出现的频率是最高的。

您好，我叫________，我从……给你打电话＊＊＊插入调查简介＊＊＊
请问您的电话号码是____________？［核查电话号码］

为了随机地从您家中选取一位我们能进行调查的人，首先需要列出您家中所有年龄在 18 岁以上的人。您只需告诉我这些人彼此的关系，不用告诉我他们的姓名。

［在将所有的人登录在表中的第 1 列之后，除非单位内只有一个成年人，否则便应继续登录他们的年龄，并提第一个问题：］我现在需要知道他们每一个人的年龄。

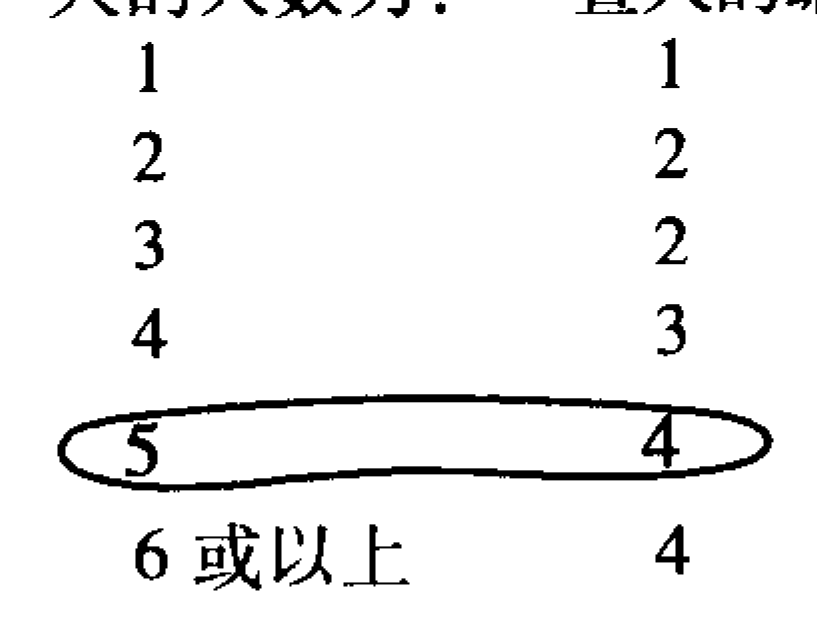
［在将所有人的年龄登录在下表的第 3 列以后，再继续问：］我这里有一张被调查人选择表，请您费一点时间帮助我们按规定选出您户中的被调查人。

列出居住单位内所有年龄在 18 岁或以上的人的年龄

与户主关系 (1)	性别 (2)	年龄 (3)	成人 (4)	记号 (5)
丈夫	男	52	2	
妻子	女	50	4	✓
女儿	女	23	5	
儿子	男	19	3	
丈父	男	78	1	

按下面所示的顺序，给 18 岁或以上的人编号：
首先为年龄最大的男性，然后为年龄次大的男性，以此类推，直至编完所有男性，然后再接着按年龄大小依次为女性编号。全部编好以后，再用选择表选取被调查人。

D 型选择表



如果户中成年人的人数为：	那么对应的被调查人的编号为：
1	1
2	2
3	2
4	3
5	4
6 或以上	4

选择图表明，我应当调查您家中的____________________________。
我是否可以和他（她）谈一谈？

［如果正在和您交谈的是另外一个人，而非您所选定的那一个人，请要求和选定的人进行交谈，并在开始提出第一个问题之前，重复调查简介。］

［如果选定的人不在现场，了解一下什么时候再打电话最合适，并将这一信息登录在呼叫单上。］

图 4.3　凯思选择表示例

调查员在确定所有合乎条件的人已经登录,并且他们的年龄都已经清楚并记录在案之后,应稍作停顿,以核对所有已经登录的人的年龄是否都合乎要求,如发现有任何不合格的人,应将他们从表中去除,并不再对他们作进一步的考虑。在完成所有这些工作之后,调查员应对所有合格的人,分性别按年龄排序编号,具体规则是:年龄最大的男性编号为1(如果存在年龄最大的男性),次大的男性为2,以此类推,直至将所有的男性全部编完并记录在表中之后,接着便是年龄最大的女性、年龄次大的女性等,直至编完全部女性,并记录在表中。例如我们考虑一个由丈夫、妻子和他们的成年儿子和女儿及丈夫的父亲组成的家庭(如图4.3所示)。这时丈夫的父亲编号为1,丈夫为2,儿子为3,妻子为4,而女儿为5。(性别的先后次序可以调换,如果研究者乐意这样做,这一选择完全是主观的。)

令那些使用凯思法的人庆幸的是,90%以上的美国家庭,合乎条件的人口都在4个以下,因此调查员在了解住户内合乎要求的人数、弄清楚他们的年龄,最后再确定并登记他们的入选编号所需的时间相对较少。在进行个别访谈时,户内的被调查人可以看到调查员在做什么,但在电话调查中却看不到。所以在将凯思法用于电话调查时,为保险起见应令调查员向被调查者作一个如图4.3所示的简要介绍,以在对合格的被调查人进行编号排序之前告诉对方这样做的原因和方法。

在登录完成了所有合乎要求的人的被选编号之后,调查员要从打印在简介/选择单上的8类选择表中选取其中的一种(见图4.3)。这张选择表将显示在所有的合乎条件的候选人中,究竟是哪一位将被指定作为这一特定呼叫单的(即抽样单位的)被调查人。在图4.3所示的选择表中,调查员接通了一个含5个合格的候选人的住户,其中最终被指定作为被调查人的是妻子(#4)。

在使用凯思法选取抽样单位内的被调查人时,应将表4.1列出的选择方面的信息资料,打印成8种类型的选择表。表4.1中的第一列显示每种表格在调查中使用的比例。在用人工进行简介/选择单和呼叫单配对时,配对的次序应为:A,A,B1,B2,C,C,D,D,E1,E2,F,F。在CATI调查中,简介/选择单由软件控制,8种类型

的选择表将按事先确定的比例随机分配给某一户。

正如我们所知,在使用凯思法选择被调查人时,需要在电话刚接通不久,就要求对方提供户内成员的具体信息。尽管这样做的确会有一定困难(参见 Czaja,Blair,& Sebestik,1982),但一个优秀的调查员完全可以做到这一点。而那些缺乏训练的调查员,尤其是没有经验的调查员,不仅会有非常高的拒访率(使无回答误差因此而有所增加),而且还可能会不恰当地使用这一方法(使未覆盖误差因此而有所增加)。

表 4.1 凯思选择表

选择表的分配比例	表编号	如果户中人数为: 1	2	3	4	5	6 或以上
		选中的成年人编号					
1/6	A	1	1	1	1	1	1
1/12	B1	1	1	1	1	2	2
1/12	B2	1	1	1	2	2	2
1/6	C	1	1	2	2	3	3
1/6	D	1	2	2	4	4	4
1/12	E1	1	2	3	3	3	5
1/12	E2	1	2	3	5	5	5
1/6	F	1	2	3	5	5	6

资料来源:凯思(Kaish)1965

总之,尽管在凯思法使用得当时,单位内未覆盖误差发生的可能性的确会有所减少。但是与此同时,无回答率却可能会有所上升,从而使无回答误差会因此而有所增加。正因为如此,研究者必须在对可能得到的好处是否会为总调查误差和调查费用的可能增加所抵消这一问题做了全面而深入的考虑之后,方能就究竟是否采用凯思法进行被调查人选择做出抉择。

被调查人生日选择法

在过去的十年间,人们探讨了用于抽样单位内的被调查人的随机选择的各种方法(Oldendick,Sorenson,Tuchfarber,& Bishop,1985;O'Rourke & Blair,1983;Salmon & Nichols,1983)。这些方法或问户

内的合格者哪一位最近过生日,或问哪一位接着要过生日。因为这些方法不仅不具任何强制性,而且能在抽样单位内产生人口特征各异的被调查人,因而受到了学术机构、公共和私营部门的调查研究人员的普遍欢迎。从理论上讲,这些方法可以产生户内的真概样本,但实际情况并非如此(参见 Groves & Lyberg,1988,p.208)。

因为生日选择法不仅既无强制性又节省时间,而且易于为调查员所掌握使用,所以非常有吸引力。调查员在作了简短的开场白之后,可按如下方式向对方提问:"我们希望调查您家中 18 岁或 18 岁以上,最后一个[即最近]过生日的那个人。"调查员必须解释最后一个的意思就是最近,方括弧中的粗体字就是给调查员的提示。在 PAPI 电话调查中,生日法还有另外一个优点,它只需要在每一个调查站张贴一份简介/选择单就可以了,而无须像凯思法那样在每一份呼叫单上都需附上一份单独的选择表。

因为这种选择方法的措辞多少有一点特别,因此户中接听电话的人也许不能马上理解调查员所说的话的确切含义,有些问题便可能因此而产生。在采用这种选择方法时,很重要的一点是调查员必须准确地向对方传达"最后一个过生日"(或下一个过生日)这一句话的确切含义。此外,生日法无须以任何形式列举抽样单位内所有合乎条件的人,因而他不能很好地控制因对方未能精确地回答我们提出的问题,而造成的被调查人的自我选择(因此,有可能产生偏倚)。

我们可以证明,在大多数情况中,但并非全部,生日法可以帮助我们从合乎条件的人中正确地选出被调查人。表 4.2 列出了 NUSL 进行的用最近生日选出的两个不同的调查中的被调查人出生月份的比例。这两个调查各有样本 1 000 个。1990 年芝加哥大都市地区调查于四月底到六月初进行,而 1992 年的全美调查则在三月中旬到四月底进行。每一个调查,调查人报告的出生月份比例最高的都是紧靠实地调查进行的那个月的前一个月,然后随着出生月份距调查时间越远而呈下降的趋势(与出生的月份的第一个英文字母无关),生日距调查时间越远,被调查人的比例越低。如果生日法运用恰当,被调查者的出生月份的分布便会与此类似。

在 1992 年的调查中,我们也向被调查人询问了家中其余的合

格者(年龄在18岁或以上者)的出生月份。在将它们和被调查者报告的出生月份进行比较之后,我们发现有20%～25%的住户,其中有生日更接近调查时间的合格者(Lavrakas, Merkle, & Bauman, 1993)。换言之,这个"错误"的人,显然乐于成为被调查人。而在某种意义上讲,类似这样的被调查人的错选,一般都发生在用生日法选择被调查人时。问题在于,至今我们仍然不十分清楚,这样的错误是否会给单位内的覆盖误差(coverage error)带来什么样的影响,如果有影响的话,又是一种什么样的影响。调查研究者十分希望以后的有关文献能在这方面有所作为,能对生日法的效度问题给出确切的解释,因为尽管它十分易于使用,但它对总调查误差产生的影响可能会超过它给我们带来的好处。

表4.2 两个RDD调查中,用出生月份与调查时间最接近的方法选出的被调查人出生月份分布

出生月份	出生月份百分比	
	1990年调查	1992年调查
一月	10.6	10.0
二月	7.9	10.1
三月	12.7	11.1
四月	12.7	7.8
五月	7.9	6.0
六月	4.5	6.4
七月	5.3	7.8
八月	6.4	8.0
九月	8.7	8.0
十月	8.5	7.0
十一月	6.7	7.8
十二月	8.0	10.0

资料来源:西北大学调查实验室

注:1990年调查于四月下旬至同年五月,在芝加哥大都市地区进行;1992年全美调查,于三月下旬至同年四月进行。

系统配额抽样法

本书前一版曾经详细介绍了三种用于抽样单位内(如住户内的以性别和年龄分层的)被调查人的系统选择法(参见Lavrakas,

1987,pp. 89~96)。我们可用这些方法得到户内的配额样本。然而自20世纪80年代中期以来,这些方法已经逐渐被各种生日选择法所取代。考虑到实际调查中发生的这样一种广泛的转变,本书只打算对配额选择法作一些简单的介绍。

特罗达尔和卡特(Troldahl & Carter,1964)提出了一种被调查人系统(并非随机)法。与凯思法相比,这种方法的强制性可以说是比较小的,但与生日选择法相比,它的强制性还是太大。布赖恩特(Bryant,1975)提出了一种修正的特罗达尔-卡特法,试图对特罗达尔-卡特法可能产生的男性比例不足问题做一些补偿。正如我们在这里所介绍的那样,特罗达尔-卡特-布赖恩特(T-C-B)选择法,是斯查佳、布莱尔和西贝斯迪克(Czaja,Blair & Sebestik ,1982)等人发明的方法的改进和发展。

T-C-B 法要求调查员提两个问题,这两个问题应作为简介/选择单的一部分:(1)"有多少年龄在________或以上的人住在您家中?"和(2)"其中有几位是女性?"第一个问题中年龄后面的下划线中的内容,取决于每一个特定的调查的需要,但在大多数调查中它都等于18(即选择成年人)。

T-C-B 法与凯思选择法有许多相似之处,首先我们必须先确定调查员应使用选择矩阵的若干个版本中的哪一个,然后再根据对方给出的两个问题的答案,客观地选择规定的被调查人。其次,在使用 T-C-B 法选择被调查人时,我们也必须将一张张单独的简介/选择单附在每一张呼叫单上(或在 CATI 中用软件相配)。在理论上讲,一个调查只要将各种版本的选择矩阵混合在一起,便能最终得到在性别和年龄分布上比较平衡(合乎配额)的,与目标总体基本一致的样本。然而,在实际调查中,即使不同版本的选择矩阵的分配和组合完全按规定进行,但最终得到的结果常常也可能是男性的比例有所不足。因此我认为,抽样库控制者应该每天都对男女性人数进行计数,并与已知的目标总体的比率进行比较,一旦发现二者之间存在差异,便可有选择地(更多或更少地)引进附有某些类选择矩阵的呼叫单,以纠正已经发现的差异,保持样本人口特征的平衡。

考虑到如何减少对被调查人的强制性,特罗达尔和卡特(Trol-

dahl & Carter,1964)提出了一种新选择法,它比凯思法对被调查人产生的强制性小。而出于同样的考虑,哈根和科利尔(Hagen & Collier,1982)又提出了另一种选择法,这种方法的强制性甚至比 T-C-B 法还要小。在使用这种方法时,调查员在完成了调查简介之后,并不要求对方提供户内年龄在某种标准以上的人口数,而是直截了当地从四种类型的被调查人:年纪最小的女性、年纪最大的女性、年纪最小的男性或年纪最大的男性中指定一种作为被调查人。与 T-C-B 法一样,哈根-科利尔法也要求在调查过程中四种不同版本的使用比例,必须与目标总体的性别和年龄比例相同或相似。在大多数情况下,如果我们等比例地使用四种版本,那么最终有可能导致样本中男性比例的不足。因此它需要在调查过程中适当增加能指定男性作为被调查人的简介/选择单的使用。与 T-C-B 法相同,哈根-科利尔法也会丢失同性别的三人或三人以上的住户中的"中年"合格者。

调查名称

电话号码

####-####

您好,我叫____________,是西北大学的学生,我们正在作一个非常简短的有关公众对心脏移植问题的看法的调查。

我是不是可以核对一下您的电话号码? ________[核对电话号码]

按调查规定,我们希望与您家中最年轻的女性交谈一下,如果您家中有女性的话。

[如果对方就是最年轻的女性,便可直接问第一个问题]

[如果她不在现场,则询问并确定什么时候再打电话来最合适]

[如果户中没有女性,那么就要求调查最年轻的男性——如果他不在场,则询问并确定什么时候再打电话最合适。]

[如果选定的被调查人来接电话了且需要的话,重复一遍调查简介,然后开始问第一个问题]

	处理结果	调查员编号
第一次试呼:	____________________	________
第二次试呼:	____________________	________
第三次试呼:	____________________	________

图 4.4　修正的 T-C-B 法呼叫单示例

在某些场合,对那些只进行一个或为数不多几个调查单元(如作为时间很短的课程作业的一部分)的电话调查的,并且没有调查经验的调查员来讲,即使使用哈根-科利尔法似乎也过于复杂了。为了能给那些缺乏经验的调查员提供一种既比较容易操作,又没有完全放弃对被调查人选择的控制的选择方法,我建议采用只有两种版本的合简介/选择单和呼叫单为一体的通用的选择方法。如图 4.4 所示,在遇到只有一个男性或女性的住户时,调查员不管是调查男性还是调查女性,遇到什么性别就调查什么性别(如同使用 T-C-B 法那样)。而在遇到两个以上同性别成员的住户时,则可要求调查最年轻的男性或女性。这种选择方法有两种版本:在户中至少有一个以上的女性时,选择女性;如果没有就选择男性。在将这种方法用于只延续几天,因而只能作为数不多的回呼的调查时,采用这样的选择方法将对抽样结果产生一些积极的影响,使青年男女有更多的机会入选样本,因为这些人(尤其是青年男性)一般很难在一个短期的调查中与他们取得联系,如果我们未能让这些人有更多的机会进入我们的样本,势必导致在最终样本中这些人的比例过低。与完全采用 T-C-B 法类似,我们建议在采用修正的 T-C-B 法时,也必须每天核对选择男性被调查人的呼叫单的使用比例。

这几种方法中,没有一种可产生一个户内概率样本。它们都是在人们普遍使用生日选择法之前,为操作上的方便而发明的方法。现在它们也许已经风光不再。

其他的被调查人选择法

在一个电话调查只需要某一种类型的调查对象时(如大学毕业的 30~59 岁之间的女性),那么就有必要使用其他的被调查人选择(筛选)法。有些调查只需要调查户主或纳税人。另一些调查也许只需要调查生活在一个比较小的地理范围中的某些人。当然也会有一些调查,要从一个一般总体中选择具有某些独特之处的人的子样本进行调查。

选择户主

自 20 世纪 60 年代以来,在美国社会,户主的定义一直都在变

化。户主的传统定义是丈夫,丈夫只要仍然健在,便是当然的户主。在20世纪50年代之前,每四户美国家庭,就有三户的户主是丈夫。而在现在的美国社会,普遍为人们所接受的更为平等的户主的定义是住户内已婚或同居的男女中主要的那一对。在只有一个成年人的住户中,户主就是那个成年人。如果一个调查需要调查纳税人(例如一个市长在作施政演说之前,希望了解他的选民的观点),那么选择户主作为调查对象,就会得到一个比较恰当的户内样本。

调查名称

您好,我叫__________,我从……给您打电话。

我可以核对一下您的电话号码吗__________?[核对电话号码]

为了系统地从您家中选出恰当的人作为我们的调查对象,您是否可以告诉我您家中18岁及以上的成年人有几个?

__________18岁及以上的成年人。

[人数为"1",如果需要,接着进行调查简介,否则便可直接跳到第一个问题]

不用麻烦您告诉我他们的姓名,但您能否告诉我他们彼此之间的关系?

[依次登录户中的成年人(如"丈夫"、"妻子"、"公公"等)]

(1)__________

(2)__________

(3)__________

(4)__________

(5)__________

(6)__________

* * 调查丈夫/当家的那对夫妻中的男性 * *

[如果户中当家的那对夫妻仍然不清楚,则进而询问谁在家庭经济中占主导地位]

请问您家中家庭的开销主要由谁承担?

[给被调查人做记号,然后继续问]

我们的调查需要和__________交谈,我是否可以和他(或她)交谈?

[如果指定的被调查人未在现场,则询问什么时候再打电话最合适,并将它记录在呼叫单上]

图4.5 户主选择示例

如果研究者希望保持户主的性别平衡,那么就可以采用如图4.5所示的选择方法。这一方法与凯思法颇为相似,调查员也从计数户内所有的成年人人数和确定他们彼此之间的关系开始。在大多数情况下,在此之后,谁是户主这一问题就比较好解决了。如果在此之后,调查员对这一问题还是不清楚,那么他接着就应该确定哪一个人,或哪一对男女在经济上处于主导地位。只要这一问题确定了,究竟是某一个人,还是某一对男女是该户的户主的问题也就迎刃而解了。

将这一选择方法的两种版本打印出来,并将它们与呼叫单匹配。其中一种版本如图4.5所示,在户中有一对居主导地位的男女时,调查男户主;而另一版本则调查女户主。NUSL在PAPI调查中采用这一方法时,将这两种版本分别打印在粉色和蓝色的卡片上,以进一步提醒调查员,在遇到户中存在处于主导地位的男女时,应该指定哪一个性别的户主作为被调查人。在只有一位户主的户型中,只要直接指定户主作为被调查人就可以了,而不必考虑卡片的颜色。

一项调查在选择户主时,样本中女性占60%或更多是不足为奇的,尤其在城市地区,妇女是许多单亲家庭的户主。我们应该事先掌握抽样区内的人口模式,以便抽样库监控人员根据每天计数的结果,随时调整不同版本的简介/选择单的使用比例,使性别比例符合计划(或首选的)要求。

筛选可能的选民

在许多与政治问题有关的调查中,其中有一些调查研究者也许只想对"可能的选民"进行调查。在这样的调查中,我们必须设计一套筛选顺序,以确定合乎要求的被调查人。尽管有关民意测验如何确定可能的选民的文章和报道可谓汗牛充栋,但至今仍无权威性的文献问世(见Lewis,1991)。而我自己使用的方法则与许多其他调查研究者可能使用的方法并无二致:要求调查员在做完调查简介之后,便开始进行被调查人的筛选。首先要求调查员用我们事先设计好的问题,甄别对方是否居住在即将举行选举的选区内。然后接着就问他/她目前是否已经在选区登记投票。接着再问那位可能的被调查人在上次选举时是否投票(假定该人那时

已经达到选民的年龄),然后再了解该人在即将举行的选举中投票的可能性。如果在筛选进行过程中的任何时候,该人表示出对投票持有消极态度(如“上次我没有时间去投票”或“下一次选举我们不想去投票”),我们便可礼貌地终止与该人的谈话(即不再进行调查)。这时我们即可令调查员在抽样单位内尝试寻找另一个可能的合格的被调查人,也可令调查员处理另一张呼叫单。

地理位置筛选

有时在一个电话调查中,电话可能会打到位于抽样区域之外的住户,因为电话的前置号,并非总是与抽样区域吻合。在抽样区域和选区相互吻合时,地理位置的筛选是简单明了的:调查员在作了开场白之后,便立即可以问:“您居住在____________?”,下划线上包含的内容就是已经确认的地理区域名称(如库克县、波士顿、韦恩县、圣乔斯和戴德县等)。如果对方告诉调查员说,他们居住的地方并不在这一特定区域时,调查员便应终止进行调查,并在挂断电话之前,礼貌地告诉对方为什么调查不能再继续进行。如果对方不能确定他们住的地方是否位于抽样区域内,调查员应该要求与户内另一位知情者交谈。

所有被排除在调查之外的住户,其所在地都应当在规定的区域之外。对于这样的呼叫单我们必须规定一种特定的结果处理码,以便抽样库监控人员对有这一类编码的呼叫单进行核查,了解这些地理位置不符合要求的住户是否比预期的更多。如果发现对方提供的有关地理位置筛选答案的精确性存在任何不确定性,我们便应对一定比例的这类编码,用反向电话簿进行核对,以有充分的理由确定是否有必要对所有这些电话号码作回访。

在抽样区域既不符合电话前置号边界,也不符合行政区划的边界时,不仅可能会给调查带来更大的麻烦,而且还常常会导致RDD电话调查根本无法进行,因为无回答误差和未覆盖误差可能会因此而同时增大。在任何调查中,无回答误差都可能由于要求一个潜在的被调查人,在刚开始接电话就立即回答有关他们居住地的详细情况而有所增加(参见Lavrakas,1987,pp.71~74)。其原因也许在于这种做法往往会使人们觉得他们的隐私权受到了侵犯,和/或使人们在被他人“探究”自己的居住地的详情(如“请问您

居住在南47街的东面还是西面")时感到的窘困或厌烦。而就未覆盖误差而言,地理位置筛选则既有可能导致删除(omission)误差,也有可能导致冗余(commission)误差。前者发生在居住在目标区域内,却回答说没有时(伪负),而后者则发生在本来不居住在目标区域内,却非说居住在该区域内时(伪正)。

在一个调查碰到这些问题时,我们应该事先对用于地理位置筛选的方法进行检验,以确定它的有效性。如果我们用抽样区域内的反向电话簿生成抽样库,那么我们也应事先对它的有效性进行确认。否则我们从正确地将电话打入恰当的地理区域内得到的好处,可能会因只使用那些已经登录的电话号码而导致的各种形式的未覆盖误差所抵消。

其他有关因素的选择/筛选

如果确有需要,电话调查的被调查对象的选择可以同时取决于几个特征。只要设计严谨,加之严格的试调查,我们就可以设计出一种既不会明显增加未覆盖误差,也不会明显增加无回答误差的可操作的选择顺序。

在选择的标准比较复杂的调查中,电话接通后,我们应该尽快询问一些用于判断对方是否是一个真正合格的被调查人的问题,以决定调查是应该继续进行,还是马上停止。否则,我们常常会因为没有这样做,花费了很多时间对一个不符合要求的被调查者进行调查,而增加了调查的费用。如果在一开始与对方接触的时候,我们就能明确而完整地告诉对方选择被调查人的标准,这种情况就可以避免。但如果我们这么做了,也许又会产生另一个问题,那就是在调查员和被调查人之间还没有建立起良好的关系之前,因为询问了对方过多的个人问题,而导致被调查者拒访。

例如图4.6展示了NUSL发明的,一种用于挑选户内任何年龄在24岁以上、非阿拉伯或犹太裔的黑人或白人男性的一组有序的选择问题。这一组有序的选择题,要求调查员首先与一位年龄在25岁以上的男性对话。如果户内没有这样的人,交谈便应该有礼貌地终止。一旦调查员与一位年龄恰当的男性取得了联系,接着应提出另一个甄别种族、民族和宗教的问题。如果在这一组有序问题中的任何一个问题上可以确定对方是不合格的,那么调查便

应有礼貌地结束。

您好，我叫__________，我在西北大学调查实验室（在埃弗斯顿）给您打电话。我们正在进行一项有关国际问题和中东问题的民意调查。

我是否可以核对一下您的电话号码__________？[核对电话号码]

本次调查的调查对象是您家中年龄在25岁以上的男士，不知您家中是否有这样的人？

[如果告知家中没有男士，有礼貌地结束谈话]

[如果告知家中没有25岁以上的男士，有礼貌地结束谈话]

[如果告知家中合格的人选当时不在家，确定再次联系的时间]

[如果对方恰好是合格的调查对象，调查继续进行，如果必要，重复调查简介：]

除了国际事务和中东问题，通过调查我们也想了解一些有关您本人的问题。现在我先来问三个有关您个人的一些基本情况方面的问题：

一、你的种族背景是：

亚裔	1[如果是，有礼貌地结束调查]
黑人	2
西班牙裔	3[如果是，有礼貌地结束调查]
白人	4
其他	5[如果是，有礼貌地结束调查]
拒访	8[如果是，有礼貌地结束调查]

二、您是否有阿拉伯或犹太血统？

是	1[如果是，有礼貌地结束调查]
不是	2
拒访	8[如果是，有礼貌地结束调查]

三、您的宗教信仰是？

天主教	1[如果是，有礼貌地结束调查]
新教	2
犹太教	3[如果是，有礼貌地结束调查]
其他，具体说明__________	4[如果是，有礼貌地结束调查]
拒访	8[如果是，有礼貌地结束调查]

图4.6 筛选-选择单示例

人们一般更喜欢把这样的选择/筛选顺序并成一句话，在开场白之后，作如下的陈述："本次调查希望调查您家中任何一位年龄在24岁以上、非阿拉伯或犹太裔的男士，不知您家是不是有这样的男士?"。以这样一种形式来表达选择的顺序，不仅常常会引起误解，而且调查员必需要从头到尾进行三次筛选，方能完全确定对方是否是一个合格的被调查人。

即使标准更为复杂的选择顺序也可以有效地用于电话调查。通过试调查，我们可以改进问题的措词和编排顺序，以避免总调查误差的加大。总而言之，使用规定的选择方法的调查员水平的高低，在很大程度上决定了在刚开始与对方接触时，因拒访或对选择方法的误解造成可能的被调查人丢失的多少。

事先与被调查人进行联系

作为选择被调查人和保证他们接受电话调查所作的工作的一个组成部分，研究人员有时需要考虑设计一种事先与被调查人进行联系的表格(参见 Dillman 等，1976；Frey，1989，pp. 127～130)。

我们之所以要用这么多的时间和精力(花费更多的费用)对可能的调查对象进行预热，是因为如果一个事先经调查的安民告示"预热"的可能的调查对象，较之未经预热，只接到一个"冷冰冰"的电话的可能的调查对象，最终更有希望成为真正的调查对象。特罗高特等人(Traugott 等，1987)的有关实验证明，事先与可能的被调查人进行联系，可使回答率提高10个百分点。花费更多的资源进行事先联系的最终目的是减少无回答误差发生的可能性，在有些情况下，则是减少未覆盖误差发生的可能性，从而减少这些误差对总调查误差可能产生的影响。

如果电话调查使用的抽样框中包含那些可能的被调查人的住址，那么事先给抽样单位或被调查人寄一封信去，是一件最简单不过的事。即使我们使用 RDD 调查，只要资源足够，我们仍然可以这样做，因为我们有可能买到商业性的地址清单或数据库(或反向电话簿)，然后我们就可以在大多数抽样区域内的抽样库中找到大

部分，但不是全部的电话号码的住址。

NUSL 的经验告诉我们，在精英目标总体，如执行总裁、律师、医生等总体的调查中，事先联系尤其有效。在这样的调查中，我们可以采用一种二阶调查法，首先给被调查对象的办公室打电话，了解他们的电传号码，然后我们再通过电传发去一封信函进行事先联系。NUSL 在过去的 5 年中进行过若干个精英总体的电话调查，使用了某种形式的事先联系使回答率达到了 60% ~80%。（注意：在所有这些调查中，都有专门的预算支付大量重呼的费用，且调查都安排在被调查人比较方便的时间。）

用作事先联系信件的格式和内容也应该仔细斟酌，一般应在抬头上标明调查的合法身份，文字长度最好不要超过一页（参见 Frey，1989，p. 129）。信中应对调查目的、内容、长度和资助者的情况做一简短的介绍。此外，在表示希望再一次与被调查人联系的同时，不要忘记告诉被调查人，我们将会对他们提供的任何资料保密。调查研究者应当以适当的方式，如在所有的信件上用非黑色的墨水签名，并在信封上贴上邮票，而非邮资已付的邮戳（假定采用邮寄，而非电传），使信件显得更为正式和郑重。我们希望大家记住的最后一点是，务必将用作事先联系信函的内容，作为一个附件编入调查员将要使用的调查简介和回撤陈述中。

在某些电话调查中，事先联系的作用不仅仅限于提醒被调查者将要对他进行一次调查。例如，在一个比较复杂，或需要被调查人在调查进行之前，从个人文档中搜集有关信息资料的调查中，我们需要事先给被调查人提供必要的指导和/或直观的帮助。事先联系必定会花费一定的时间，并增加一定的费用，所以调查研究者还必须事先就经费的预算和分配做出决定。

练　习

1）为 RDD 调查写一篇调查简介。该调查的样本为某地方行政辖区（县、市或镇）中的成年男性，调查目的是了解他们对市政府提供的服务的满意程度。假定该调查是受地方政府部门（如县议

会、市长办公室或镇政务院)的委托。

2)用习题1)编写的调查简介,设计一张采用本章介绍的任何一种在每一样本住户中选取一个成年人的有控制选择方法的简介/选择单。

3)继续上面两个习题,写一篇调查员用以回答可能的被调查人提出的这样一个问题:如果他们参加市政府委托进行的调查,将如何为他们提供资料保密的回撤陈述,并注意使它对那些认为他们的名字是由市政府提供的可能的被调查人更具针对性。陈述的文字不超过100个字。

4)编写一封能寄给每一个可能的被调查人的用作事先联系的信件。被调查人的地址来自上面的调查。

5)联系当地一家市场调查事务所,并了解他们常用于一般人口总体调查的被调查人选择的方法,以及他们之所以采用这些方法的原因。写一份调查报告,说明你的发现,如果有必要,报告也可以包括你对事务所选择的方法的意见。

6)为一个测量购物习惯的全国性的RDD调查撰写和编排一份调查简介和被调查人选择顺序。其目的在于筛选如下类型的合格者:男性、年龄在25~44岁,至少要有一定的在研究生院学习的经历。

7)对习题6)中设计的选择顺序进行适当的修改,以排除所有那些没有American Express card,Master-Card或Visa card的人。



督导Ⅰ:安排调查员的工作

Supervision I:Structuring Interviewers's Work

电话调查和个别访谈中的质量控制问题
如何招聘调查员
怎样进行培训
调查前实习和在岗培训
总调查误差和调查员
练 习

诚如格罗夫斯(Groves,1989)所言:"在[个人访谈和电话]调查中,调查员是测量得以进行的中介"(p.404)。这就是说,调查员的工作不仅仅是提问和记录,而且还要处理样本和寻求被调查人的合作。如果有人说调查员的工作是至关重要的,那也不足为奇。因为他们可以在不经意间给调查量度增加显著的偏倚和变差。然而尽管有许多能减少与调查员有关的误差的策略和方法可资利用(参见 Fowler & Mangione,1990,p.9),但人们常常弃之不用。

访谈是整个电话调查过程的一个组成部分,与其说它是一种科学,不如说它是一种手艺。在这一章我们将给大家介绍调查员在电话调查中的重要性。主要目的是想使读者了解究竟怎样才能得到即使不能使调查数据中的测量误差、无回答误差或未覆盖误差完全避免(实际上也不可能完全避免),但能降到最低的优秀的调查员。为了达到这一目的,我们必须精心计划和安排调查过程中从调查员的招聘到在岗培训的一系列工作。

在这一章,首先比较电话调查和个别访谈之间在访谈质量控制问题上存在的差异。然后,再介绍如何招聘有酬和无酬(志愿)的电话调查员。其中也会谈及在最后决定聘用什么人作调查员之前,如何对应聘者进行甄别和筛选。

无论是 PAPI 还是 RDD,每次调查正式开始之前,对所有的调查员进行培训都是很重要的。在这一章我们将分一般总体调查和特定总体调查两种不同的调查,分别对它进行详细的介绍。访谈用语训练包括访谈时如何正确无误地表达问卷中的封闭和开放式题项,以及如何无偏地对那些模糊不清或不完整的答案进行追问。此外本章也对调查员的训练和如何降低无回答误差分别进行介绍。随后,我们又对被人戏称为"永远也结束不了的"调查员的在职培训问题做了介绍。在本章的结论部分,我们就有关调查员对总调查误差可能产生的影响,以及我们可以用于估计这类影响大小的方法展开一些讨论和介绍。

电话调查和个别访谈中的质量控制问题

本书对如下的观点予以特别的关注:与其他各种数据搜集法相比,电话调查可能存在一个非常大的优点,即它能使数据搜集过程变得更为集中,从而给我们提供更多质量控制的机会。在电话调查中,如果质量控制方法使用得当,那么它最后得到的数据必定是高质量的,因为更为集中的数据搜集过程在相当大的程度上减少了调查员在调查过程可能发生的各种错误行为。

令人不解的是,虽然许多调查研究者似乎都已经认识到有代表性的抽样库、低无回答率和设计严谨的问卷的重要性,但往往疏于对电话调查的实际过程进行控制。这也许是因为在这些人看来,只要问卷设计得好且抽样方法得当,调查数据的搜集就自然会成功。而就用电话调查搜集调查数据的实际过程和就它可能对总调查误差产生的各种影响看,这种认识显然是不明智的。除非数据搜集的过程是以某种可控制的和标准化的方式进行的,否则调查搜集到的数据将可能是无效的(参见 Fowler & Mangione,1990),因而也是徒劳无功的。这样的调查无疑是对资源的极大浪费。

缺乏对电话调查访谈过程的严格控制的原因很多,但主要是因为经费不足。尽管对电话调查访谈过程进行严格而一贯的控制的费用的确不低,但我们必须懂得,如果没有这样的控制系统,那么调查机构花在其他方面的钱(如抽样)也许会血本无归。

不重视电话调查访谈过程的控制的另外一个原因可能是,许多调查研究人员通过传统的面对面的访谈方法来学习电话调查方法。在用个别访谈搜集数据的实际操作过程中,调查员和督导员的作用与电话调查存在较大的差别。例如,在面对面的调查中,我们无法立即对一份已经完成的问卷进行检查,并将检查结果立即反馈给调查员。不仅如此,与电话调查不同,个别访谈的督导员也无法对任何正在进行中的访谈进行监控。

此外,因为其他的各种原因,较之电话调查,个别访谈的调查员在搜集数据过程中对数据质量的影响更大。无论调查员的水平

如何高，面对面访谈得到的数据，尤其是在标准化程度上，都无法与用高质量的集中化的电话调查搜集到的数据可能达到和可望达到的水平相比。

如何招聘调查员

普遍需要考虑的问题

调查员招聘中所应考虑的几个基本问题，既适用于有酬的调查员招聘，也适用于无酬的调查员如志愿者或将调查作为课程作业的一部分的学生的招聘。相对而言，较之招聘无酬的调查员，招聘有酬的调查员能招聘到高水平的调查员的可能性更大，其原因大致有以下几个：首先，因为招聘的调查员是有酬的，所以挑选的余地比较大。我们可以从较多的应聘者中精心挑选其中水平最高的那些人。其次，有酬的调查员更可能与调查目的保持一种客观的超然的态度。相反，那些无酬的调查员常常会对调查资料有所图。这就是说，志愿者本身的性质决定他们在进行一项调查时常常负有组织交付的使命，对调查结果有所预期，而这就会对他们作为一个调查员的行为发生影响，从而使他们搜集的数据发生偏倚（测量误差）。无独有偶，为学术信誉而参加调查的学生，常常也会对调查结果感兴趣，如果调查就是他们本班的项目时，那么情况尤其如此。

每一个被聘的调查员，无论是有酬还是无酬的都应当和聘用他们的调查研究者签订一份书面的协议。协议条款应包括不侵犯被调查人的隐私这样的条款。虽然书面协议对聘用有酬调查员来讲，是招聘过程中所必须经历的一种标准程式，但对于那些聘用无酬调查员的人来讲，这样的情况似乎不多见。而美国志愿行动组织（NOVA）强烈推荐，在聘用志愿者做他们自己自愿承担的工作时，应把他们与有酬的雇员同等对待。NOVA 认为，让志愿者签订一份无约束的工作合同（nonbinding work contract），可以提高志愿者对他们同意承担的工作的重要性的认识。无酬的电话调查员必

须懂得,他们的时间是不会白费的,而一份书面合同对于他们更加确认自己所做的工作的重要性具有某种象征意义。

无论调查员是有酬的还是无酬的,就调查员培训问题而言并无二致,但依据调查所需的调查员的类型的不同,二者在聘用的方式上则可能会有所不同。尽管如此,在所有类型的调查员的聘用中仍然存在着某种共性。首先,我们必须对所有的调查员讲清楚,电话调查需要"标准化的访谈"(参见 Fowler & Mangione, 1990)——高度结构化和非常刻板的访谈。所谓标准化访谈是指在调查过程中,调查员既不能任意更改问卷题项的措词和前后次序,也不能自作主张地决定选择什么人作为调查对象。不仅如此,我们还应当告诉所有未来的调查员,在整个调查过程中,包括访谈过程始终都会处在督导人员的监控之中。

在签约之前告诉未来的调查员有关调查的事情,将使他们对未来的工作抱有比较实际的期待。在招聘有酬调查员时,诸如此类的要求有可能使那些不喜欢在高度结构化情境下工作的应聘者知难而退。此外,每种类型的调查员的招聘还有其他应该考虑的问题。

招聘有酬调查员时应考虑的问题

在招聘有酬调查员时,我们有机会聘用到高质量的应聘者。所谓高质量的应聘者是指那些最有可能执行那些设计用来最大限度地降低总调查误差的方法步骤的应聘者。为了达到这一目的,我们必须对招聘过程作精心的计划安排。我们应该先用精心设计的筛选程序发现有能力和有经验的人才,然后再用比较优越的报酬吸引他们,以达到我们招聘到高水平调查员的目的。因为如果报酬过低,比较有能力的人就不会对电话调查感兴趣。总之,我们付的酬劳越多,无论在数量上还是在质量上,我们将会得到的回报也越多。我们甚至可以说,就成本效益而言(就总调查误差的减少而言),雇用一个报酬较高但调查的质量和数量都比较高的高水平的调查员,其成本效益肯定高于聘用一个报酬相对较低但水平不高的调查员。

在招聘新调查员的广告中,除了说明以钟点计算的报酬的起

点之外,还应该说明工作的性质是临时性的。例如,在 NUSL,调查员每周工作 12 ~20 小时。而在一些大型的调查机构,如密西根大学的调查研究中心,调查员工作的时间更长一些(18 ~35 小时/周)。在一定程度上,这也许是他们的机构更大,接受的调查工作也更多的缘故(Groves,1989)。

我认为在招聘广告中最好加入"有经验者优先"这一条。因为这将有助于筛选掉那些缺乏经验的应聘者。在这里,有经验是指从事电话调查,而非电话推销的经验(参见 Frey,1989,p. 220)。(实际上,某些只有一定电话推销经验的人,可能会在适应电话调查要求的标准化电话访谈的与电话推销的行为方式大相径庭的行为方式时感到非常困难。)如果我们招聘的调查员是用来做 CATI 调查的,那么打字和使用计算机的经验也应当优先考虑。NUSL 的广告中给出的起始报酬约为当时最低报酬(每小时 4.25 美金)的 2 倍。付给新调查员报酬的高低,主要取决于 NUSL 希望招聘到的调查员水平的高低,其次取决于芝加哥都市地区生活费用的高低。

我认为招聘广告上应该刊登联系电话,而非接待应聘者来访的地址。用电话筛选一个未来的电话调查员的根本目的在于,这样能评估应聘者在"在电话上的行为举止"。如果未来的调查员直接来应聘,那么他的言语之外的其他举止行为可能会影响招聘者的判断。迄今为止,可视电话尚未普遍应用于电话调查,影响电话调查最终是否会取得成功的因素是应聘者的声音和其他语言行为,而非他们的外貌。因此我们没有什么必要在招聘的最初阶段面见应聘者。(但这并不是说,个别面试不能搜集到任何重要的信息。例如,在个别面试时,我们就可测试应聘者的朗读能力。)在用电话选拔调查员时,应该使用一种测试表,测试的项目除应聘者的人口学和有关背景资料之外,还应该包括招聘者从电话中得到的应聘者的"第一印象"的得分,这一得分反映应聘者通过电话表现出来的语言能力的高低。应聘者的经验越丰富,第一印象得分值就越高。

调查的组织者也许会考虑这样一些问题——是否有什么人口学特征,如性别、年龄或教育程度可能与访谈质量有关。如果有,那么随之而来的问题是,在做出是否聘用应聘者的决定时,我们是

否应对这些特征加以考虑?就力图避免在实际聘用过程中的偏见这一问题而言,应该引起我们注意的问题是:“除了良好的朗读、速写能力和[在他身上表露出来的]在情理之中的快乐的个性之外,不存在任何其他令人信服的,用以识别可能的调查员的选择标准”(Fowler & Mangione,1990,p. 140)。即使在地方口音非常重的地方,巴斯和托特拉(Bass & Tortora,1988)也没有发现任何与调查员自身的某种特征相关的在选择调查员时应加以考虑的因素。

另一方面,如果调查的题目与调查员的人口学特征有关,那么我们便有足够的证据证明,的确会有我们所不希望看到的影响调查员/被调查人而导致总调查误差升高的情况发生,而这些影响将导致总调查误差的增加(参见 Fowler & Mangione,1990,pp. 98 ~ 105)。例如,一个在本书正在写作当中进行的一项有关性骚扰问题的调查中,我们发现男性的被调查人在接受一个男性调查员调查时,报告其在工作中曾对某人有过性骚扰的可能是他们在接受女性调查员调查时的两倍(Lavrakas,1992)。在这样的情况下,我们在聘用标准中必须考虑调查的目的,并采用一视同仁的方式对调查员的人口学特征加以考虑。总之,如果我们有十分明确且令人信服的证据证明在调查员选择标准中考虑调查员的人口学特征,可以降低与调查员有关的可能影响,那么我们必须以某种无偏见的方式对待那些“不合格”的应聘者。

假定大多数电话调查聘用的调查员都是临时性的,那么聘用和培训的组织方式都应该是某种自我选择的过程。在这一过程中,应聘者将自行决定他们是否可以成为一个合格的调查员,如果不能,便自行撤回自己的申请。根据初选的结果,那些被认为确实具有合格调查员潜质的人将被选中接受培训。在培训过程中,自我选择仍在继续。例如,在对应聘者进行初试时,判断他们是否能成为合格的调查员的重要标准是可靠性(即能按时完成计划规定的访谈过程中的每一阶段的工作),而在培训阶段,我们则需要采用另一种方式对通过初试的应聘者作进一步的考核,即对某一特定的应聘者的适用性进行考核。在通过这两项考核之后,我们则应对入选的应聘者进行在岗考核,即在他们实际操作电话访谈的过程中,对他们进行最后的考核,以便决定是否继续聘用他们做调

查员(参见第6章)。

聘用无酬调查员时应注意的问题

可用于控制可资利用的不付酬的调查员的手段,比付酬调查员的要少得多。在访谈作为课堂作业的一部分时,一般班上的所有学生都必须参加资料搜集工作。这时督导员/指导员就必须使每一个注册的学生都参加调查,而通过精心的培训和现场督导,大多数学生所做的调查基本上还是可以接受的。在使用志愿者作调查员时,我们可以采用类似上述用于有酬调查员的自我选择和筛选标准。但是,在选择结束时,研究者必须主动告诉那些未入选的志愿者,他们也许更适合从事其他类型的工作,而不是电话调查。

怎样进行培训

福勒和曼乔尼(Fowler & Mangione,1986)曾对培训与调查员的关系,即培训对调查数据质量的影响进行过全面的研究,虽然这一研究发现这些方面的影响不是很大。尽管如此,它仍然足以令研究者得出这样的结论:"调查员在一般调查技巧方面需要进行有督导的练习……[培训]只用来朗读、演讲和演示一次访谈过程是远远不够的"(Fowler & Mangione,1990,pp. 117~118)。这些研究结果说明对调查员进行多方面的和分阶段的培训是十分重要的。

在一个新的调查正式开始之前进行的正式培训是调查员培训的核心。这种形式的培训目的在于使调查员(和督导员)能有机会:了解和掌握足够的背景资料;学习和掌握调查的实际操作程序;学习和掌握各种精心准备的有关在岗培训和/或实际访谈的材料。在培训阶段需要传达给调查员的信息包括两种类型:(a)适用于所有调查的通用操作规范和一般要求;(b)适用于某一调查的特定的操作规范和要求。

培训一般应由那些富有调查经验的人员主持,若能由主管实际调查的负责人主持则更好。主持人应该了解调查过程涉及的各个方面,最理想的培训主持人应该是那些既担当过电话调查的调

查员,又担当过电话调查的督导员的人。主持培训的人必须具有较好的表达能力和组织能力:在培训阶段出现事先未曾预想到的情况时,他们仍然能条理清晰地以非常得当的方式与调查员讨论商量解决的办法。总之,培训主持人在这一阶段的责任就是,使培训对象逐渐懂得培训内容对实现高质量的调查的重要性,从而为下一阶段的进一步培训或实际调查做好准备。

为了组织好培训,每一调查组织都应编写好一套调查员(督导员)培训手册。培训手册应该对那些在调查组织者看来是进行一个标准的访谈所不可或缺的基本知识和必须掌握的基本步骤进行全面而详尽的介绍。手册应包括所有在培训阶段必须进行讨论的内容(见下面的有关章节)。对这一问题有兴趣的读者,可参阅弗瑞(Frey,1989,pp. 222 ~ 233)有关电话访谈培训手册的讨论。

一般工作经验培训

下面我们要向大家介绍的问题是有关一般操作方法和经验的培训,它包括以下三个方面:

1. 一个优秀的电话调查员所必须具备的能力包括正确地处理抽样库、扼要地介绍调查、按规定选择和筛选出被调查人,并确保选出的被调查人的合作和以标准的方式进行问卷调查等;
2. 从质和量两个方面阐述如何对访谈进行监控;
3. 如何满足委托进行调查的组织或个人对调查的特殊要求。

福勒和曼乔尼(Fowler & Mangione,1990)认为,规范的电话调查员的培训时间不能低于 20 ~ 30 小时。培训组织者应对这一问题有充分的认识,并在预算中考虑培训所需的各项费用。

高质量的电话访谈

每次培训都应该详细讨论(或回顾)一个优秀的电话调查员所应展示的行为特征是什么。如果时间不允许,我们不能对这一问题展开详细的讨论,那么对它作一番简单的回顾,也将对未来的调查有帮助。读者也许不同意我在下面对这一问题提出的具体建议,但是大概都不会对人们一般都会对调查员的言谈举止有一定要求这一点持有任何异议。

无论是使用 PAPI 还是 CATI,一个优秀的电话调查员都应该以

某种积极的方式恰当地使用调查呼叫单、调查简介/被调查人选择程序、回撤陈述和问卷。他们也不能为完成一定的访谈数量而牺牲访谈质量。通过培训，我们必须使调查员懂得，为了实现高质量的电话访谈，他们要做的事，远不只是像以标准化的方式宣读问卷并记录下被调查人给出的答案那么简单。

有关调查处理结果码、调查介绍、被调查人选择顺序、回撤陈述和问卷的详细讨论我们将在培训的第二部分介绍。在第一部分的培训中，我们应该使调查员知道，调查过程中的每一个步骤都是一个完整的调查不可或缺的组成部分。不仅如此，我们还必须使调查员知道，它们对总调查误差可能产生的影响（参见 Fowler & Mangione，1990）。实际上，围绕总调查误差的概念来构建我们的一般性培训教材，正是我们一贯所推崇的。

培训教材不仅应该告诉调查员如何以令人愉快的方式与人进行交谈，而且还应告诉他们不应偏向（即强调）某些形式的回答。在被调查人超出问卷列出的特定内容和他们详细讨论自己的观点或问题时，我们应当提醒调查员不要使自己变成"电话机上的临床医生"。调查员必须明白他们必须时常提醒自己，要为保持和控制调查过程的进度和谈话方向而不懈努力（最理想的情况是，问卷的措辞和编排的方式能有助于调查员保持对调查过程的控制）。

在许多调查场合，调查员举止行为多少有点像一个聪明的机器人。他们不应当让自己的个性过于热情，或过于超脱而使被调查人给出的答案发生偏倚。调查员必须在这两种极端的态度之间保持平衡，这样便可以使被调查人在不受到任何来自调查员方面的影响的情况下，给出他们自己的答案（参见 Fowler & Mangione，1990，pp. 48 ~ 50）。

福勒和曼乔尼（Fowler & Mangione，1900）认为只有像"调查员以一种照本宣科的方式向被调查人提问，除此之外，调查员只对问题的目的作了一些非常有限的解释且逐字逐句地记下了被调查人给出的答案"（p. 68）这样的访谈，才可以称之为标准化的调查访谈。他们解释说，这样的访谈也包括"以某种无导向的方式为澄清和详细地说明有关概念所做的[探讨]"和"价值中立的、不带任何主观判断地就问题答案的实质[进行交流]"（p. 33）。

与此同时,调查员应该使被调查人明白自己正在接受另一个人而不是在接受一个录音机的调查(如同某些电话推销员进行的电话"调查"那样)。我们必须允许调查员,而且应该希望调查员培养个人的判断能力,使他们不需经过可能引起被调查人答案产生偏倚的评论,就知道如何与不同类型的被调查人打交道。

培训人员也应该了解,让调查员懂得他们的任务是与被调查人打交道这一点很重要。格罗夫斯(Groves,1989,pp.407~440)有关这一过程的认知科学方面的文献回顾,对那些计划将这一方面的知识编写进为调查员准备的一般培训手册的人可能特别有用。例如,它可以帮助调查员懂得被调查人只有在经历了一定的步骤之后,才能回答他们提出的问题。这些步骤大致是下面5个:(a)与提出的问题有关的从前的信息编码;(b)对调查问题的含义的理解;(c)记忆中可以回忆的有关知识;(d)在备选答案中作判断,并选择恰当的答案;(e)向调查员报告他们的答案。

间接询问和间接回应

在培训的第一阶段,培训人员应该向调查员说明,在被调查人以文不对题或含糊不清的语言回答问题时,他们应该如何以不会对被调查人给出的答案产生影响的方式进行间接的询问。调查员也应该了解如何以非直接的方式,向被调查人做必要的回应。查尔斯·卡内尔(Charles Cannell),一位在调查访谈的行为举止和有关调查员对调查结果的影响研究方面的学术带头人认为,在调查员就被调查人给出的初步答案进行询问(追问),或做出其他回应而与被调查人进行交谈时,主要应该关注被调查人在回答问题时的行为举止(行动),而不是从这些答案中得到的感受(Charles Cannell,1991)。例如,调查员使用的关注行动而非感受的无偏评语可以包括:"这是非常有用的信息","这正是我们需要的信息"和"您是否可以把最后那个问题再讲得具体一点?"等等。我们应在培训过程中告诉调查员,为什么诸如这样的表达方式,要优于类似"发生这样的事太糟糕了",和"不过您刚才提到的问题的确很重要"这样的把关注的重点放在自己的感受上的表达方式。福勒和曼乔尼(Fowler & Mangione 1990,pp.37~46)和弗瑞(Frey,1989,pp.224~225)在他们的有关著作中不仅对无偏询问和回应做了详

尽的介绍,而且还列举了若干使用的实例。

无论是PAPI还是CATI,我们都可以通过问卷的编排方式来帮助调查员进行无偏的询问和回应。我们可以在问卷中编入一些非直接询问和价值中立的回应评语的实例供调查员使用。在培训中我们应该告诉调查员如何正确使用这些刊印在问卷上的评语,以降低与调查员有关的各种变差。此外,对问卷中的无论是已作文字定义,还是未作文字定义的题项,其定义的方式都会对调查员对被调查人不太清楚的问题所做答复的质量产生影响。当被调查人对问卷中未经定义的题项存有疑义并要求调查员进行解释时,正确答案应该是"无论您怎样理解都可以"。

如何避免拒访

资质优秀的调查员与资质一般的调查员之间差别似乎只有一个,那就是他们在处理那些难以打交道的和直截了当地拒绝接受调查的被调查人的能力的不同。因此,在用于对调查员进行调查的一般要求的培训材料中,应该对拒访的性质进行详尽的讨论。此外,培训资料还应清楚地告诉调查员,他们应当如何不带强制性地和有礼貌地说服对方接受调查。对这些问题有兴趣的读者,可进一步阅读格罗夫斯撰写的社会科学文献中,有关说服和服从问题的评论。这些评论都与被调查人接受调查的意愿和降低无回答误差的访谈策略有关(参见Groves,1989,pp.215~216)。

与我在本书原版中所写的内容一样,我仍然认为我们最好假定所有的潜在的被调查人都需要我们给他们提供某种他们之所以应该参加调查的鼓励。就大多数被调查人来讲,只要我们能告诉他,他提供的答案将对我们有很大的帮助这样的鼓励就足够了。而对另外一些被调查人,似乎需要使他们感到他们是举足轻重的,他们的观点正是我们在寻找的。

但在一般人口总体和特定人口总体的调查中,所有潜在的被调查人中20%~30%需要调查员使出浑身解数向他们去"兜售"调查。在这些比较困难的场合,一种可供我们选择的方法是假定时间不合适,即假定潜在的被调查人之所以拒绝接受调查,是因为调查员联系进行调查的时间不合适,因此为了把调查时间改在对方认为合适的时间。我们应该使调查员学会说:"对不起,打扰了。

我们打电话的时间对您似乎不合适。"然后，调查员必须根据自己对当时情况的判断，或询问对方什么时候再给他打电话最合适，或告诉对方还会给他打电话，或不再说话。另一种方法是"恳求"潜在的被调查人。在调查问卷相对较短（譬如只需要10分钟，或不到10分钟）的情况下，调查员可以试着使不愿接受调查的被调查人相信调查不用花费很长时间。还有一种说服不愿接受调查的被调查人的办法是告诉他们可以拒绝回答任何他们不愿意回答的问题（虽然这样做可能会使无回答误差有所增加，但经验告诉我们，它更有可能使被调查者感到比较放松，因而实际上并未导致更多的数据丢失）。此外，通过培训我们还应当让调查员学会如何用调查回撤陈述，对调查的合法性和重要性等问题做出恰当的解释，并能针对不同的调查人给出不同层次的保证。有时被调查人希望对调查的背景做进一步的核实，这时如果调查员能告诉他们调查的组织资助者的姓名和电话常常会使问题迎刃而解。在这样的情况下，只是做一些简单的保证则可能于事无补，远不能消除被调查人对调查心存的疑虑。

最后一个问题是调查员应该了解被调查人的合作是在帮助他们挣钱（对无酬调查员而言，被调查人的合作是在帮助他们完成任务）。调查员必须将寻求合作的努力个人化，不要只是说一些诸如"给更好地设计一个社会项目以帮助"这样的空话，给对方以一些空洞的鼓励；也不要试图以第三方的大名（调查的资助者）来说服对方。相反，那些不愿接受调查的被调查者，也许在了解假如他们接受调查的话，将会给调查员带来直接的帮助之后，可能会改变态度接受调查。（这就好比本章后面将会提到的那样，用于支付调查员酬金的方式可能会影响调查员为减少拒访而作的努力。）

在培训中，培训人员不仅要告诉调查员在调查过程中，为了在最大程度上减少拒访的发生，他们应该使用一些什么样的语言，而且还要告诉他们说话时应该掌握的态度和使用的语气。考林斯等人（Collins 等，1988）的研究发现，那些不太成功的调查员，在遇到问题时"表现出缺乏信心和流露出一定的恐慌，他们似乎对遇到的问题缺乏准备，轻易地退却且没有设法避开'死胡同'"（p. 229）。

在电话调查的无回答率不断上升且其中拒答所占的比例最大

的情况下，我们必须对调查员进行如何避免拒答的培训。正因为如此，不断为调查方法中这个十分重要的领域提供新的文献是调查研究者和培训人员的重要职责。诚如我们在第 3 章所指出的那样，想方设法争取转变拒访者的态度已经成为了许多高质量调查的一个不可或缺的组成部分。我们只需对那些将来有可能实际从事拒访转化工作的调查员进行这方面的培训。因此它不宜在所有调查员都必须参加的一般工作经验培训阶段进行，而应该在实地调查工作马上就要开始之前（一般在第二个星期的后半段）进行。

角色扮演

角色扮演的培训也应该是一般工作经验培训内容的一个组成部分。例如，调查员可以练习如何说服不愿接受调查的被调查人接受调查，这时他们的角色就如同培训阶段的培训人员。角色扮演也是一种演示在 PAPI 问卷中如何做跳答，或学习询问各种开放式题项的有效方法。

角色扮演式培训与那些只是由培训人员滔滔不绝地宣讲、单纯收看培训教材录像带和阅读培训手册式的被动式培训不同，它能调动起调查员的积极性，变被动接受培训为主动参加培训。正因为如此，大多数调查员都乐于参加这样的培训。

CATI 式调查的培训

在 CATI 式调查中，角色扮演也可以与计算机操作的练习合在一起进行。不过调查员在接受 CATI 式的角色扮演培训之前，必须熟悉他们使用的 CATI 系统的操作方法（硬件和软件）。无论使用何种 CATI 系统，调查员都必须明白他们的"主要工作是访谈，而不是操作 CATI 系统"（House & Nicholls，1988，p. 427）。

在 CATI 式调查的调查员培训中，应该先向调查员说明 CATI 系统是如何提高访谈质量和效率的，从而使与调查员有关的误差有所降低（参见 Groves，1989，pp. 377 ~ 378；Lavrakas，1991 或参见第 1 章）。例如，CATI 系统对跳跃式（有分支）题项，即那些内容与被调查人在前面的题项中给出的具体答案有关的题项，以及那些需要在次序或措词上作随机变化的题项有很大的帮助（参见 House & Nicholls，1988）。培训必须使调查员明白，使用 CATI 系统的目的

在于提高数据的质量(即降低测量误差)。与 PAPI 调查相比,CATI 调查的费用会略高一些,产生这一现象的部分原因在于完成访谈所需的时间会稍微长一些。

由上述可知,在使用 CATI 系统进行电话调查时,只对调查员进行硬件和软件操作方法的培训是不够的。在 CATI 调查的培训中,我们还应该使调查员逐渐了解这种先进技术对提高数据质量和降低其他方面的总调查误差的潜在价值。

讲解督导系统

为了便于大家的理解,也许应该将面对面访谈和电话调查在督导员作用上存在的差异作一番比较。在这里我们特别想要强调的是,集中化的电话访谈给我们提供的对访谈过程进行的高度监控是如何在最大限度上降低了总调查误差,尤其是测量误差的。此外,我们还要请大家注意,较之面对面访谈,集中化的电话访谈由调查员引起的误差似乎也比较小。

调查员必须记住他们将要完成的那些例行工作——他们将会分配到一定数量的呼叫单,按规定使用的特定的处理结果码,必须一一核对呼叫单上的电话号码,使用规定的选择程序在每一住户中选出一位合格的被调查人等等——而督导员的责任在于监控调查员按规定完成的工作。我们也应该告诉调查员,督导员将会从许多方面来检查他们的访谈质量,例如问卷完成之后应该立即交给督导员,以便他们立即确认问卷的完成情况。不仅如此,一个理想的集中化的电话调查库,应该允许督导人员监听调查过程中的所有情况。我们应该使调查员明白,整个访谈过程都处于督导人员的监控之中,这样他们才可以随时得到来自督导人员的建设性意见,以尽可能降低总调查误差。最后,我们还应该使调查员认识到,督导人员将会从他们完成的问卷中挑选一部分进行全面的检查(参见第 6 章)。

在解释对访谈过程作深度监控的必要性时,培训人员应该使调查员明白,监控的目的在于保证数据质量,而高质量的数据则体现了他们对调查资助者负责的职业道德和职业素质。不要使调查员对监控产生误解,认为这是对他们的不信任,相反应该使他们懂得督导人员的监控是在帮助他们尽可能最好地完成访谈任务。

访谈效率

调查员应当明白，提高访谈效率同样也很重要。调查员访谈效率的一个基本量度是，他们在每一访谈阶段中完全完成的问卷数和拒访（和部分拒访）的问卷数之比。而我们的目标则是通过培训，使调查员完成的问卷数和未完成或部分完成的问卷数之比达到 4：1，从而使回答率能达到 80% 或更高。这一数字与密西根调查研究中心（Michigan's Survey Research Center）的调查员的回答率均值基本相等（Groves，1989，p. 195）。当然这一比率会随抽样区域的不同而有所不同。根据我们的经验，假如抽样区域主要是城市，那么这一比率也许定得过高。在这种情况下，调查员可把它视为自己努力的目标。我们还应该告诉调查员所谓恰当地完成了这一句话的确切含义是什么。此外我们还应该告诉他们拒访和无法联系为什么可能会在不知不觉中损害了调查的有效性，而使总调查误差有所增加。

在讨论调查员的效率问题时，培训员必须注意到每小时可以完成的问卷数将取决于问卷的长度和抽样库能使调查员联系到合格的被调查人的几率（见第 2 章）。为了便于比较，我想用自己在 RDD 调查时得到的有关经验对这一问题做一些更为具体的解释。在问卷平均长度在 25～30 分钟左右时，一个调查员每小时平均可完成调查问卷一份。而在问卷的长度在 15～20 分钟时，每小时平均可完成两份或不到两份。问卷越短，每小时完成的问卷就越多。调查员应该明白督导员对他们每小时完成的问卷数是有要求的，特别是在他们的报酬在一定程度上与完成的问卷数有关时，情况尤其如此（如以后将要讨论的那样）。

调查员也应该明白不仅他们每天的工作效率都将记录在案，而且会对他们访谈的拒访率和完成率进行比较。调查员必须认识到，如果他们总是有太多的拒访的话，他们的聘用期便可能难以为继，因为调查无回答率的增加最终将导致总调查误差的增加。

最后，我们还应该告诉调查员，他们应该以什么样的速度来处理呼叫单。调查员不应该匆匆忙忙一次接一次地进行拨号。相反他们应该了解两次拨号之间，应有某种视情况不同而长短不同的间隔。

如果用于督导的资源允许,我们可以专门设计一种监控方式来评估访谈的质量(参见第6章)。如果我们准备这样做,那么在进行一般工作经验培训时,我们也应该对这样的监控方式作介绍,以便调查员对用于评估访谈质量的标准有所了解——例如,调查员在多大程度上正确地照本宣科地宣读了问卷中的问题,调查员有多少次作了不必要的评语等。

在使调查员牢记数量的重要性的同时,也必须使他们牢记质量的重要性。毋庸置疑,较之数量,质量更为重要。但是大多数调查的预算一般都比较紧,因此不能单纯强调质量,必须根据实际情况,兼顾质量和数量,使二者保持平衡。

付酬方式和其他与调查员聘用有关的规定

培训员在进行有关一般工作惯例及那些不论是在什么人组织的调查中,都可能会有的某些特定方法和要求的培训时所应该讨论的最后几个问题是:聘用什么形式的调查员;工作如何分配;考核方针;怎样和什么时候支付调查员的工作报酬(在调查员是无酬时,怎样和何时给他们出具工作表现的证明)等等。

报酬问题非常重要,必须予以专门的关注。从心理学的角度看,理想的付酬方式应该体现我们对调查员所抱的期望。标准的用于调查员计酬的方式是按小时计酬。不考虑效率的按时计酬方法的背后隐含着这样两种观点:第一种观点认为这种方式最便于行政管理;第二种观点则认为它是质量重于数量思想的最好体现。后一种观点是建立在一种假设之上的:如果调查员以计件的(以完成的每一份问卷计算)方式付酬,那么他们将被迫尽快地完成每一份问卷,而调查质量则会因此而被牺牲。

尽管许多专门从事调查的人持有这样的观点,但我依然认为只以业绩为依据的个人化的付酬方式,忽略了为保证调查质量而建立起来的高度结构化的监控系统的作用。根据本人以往15年中在电话调查方面的经验,我认为混合式付酬方式是一种比较好的付酬方式。在采用这种付酬方式时,调查员报酬中的大部分将以标准的小时计酬方式支付,而另一小部分则以效率为依据支付。一般来讲,这两部分之比应以3:1为宜。例如,一个调查员每小时的报酬可能为6美元,那么另外平均每小时2美元左右的报酬,则

应以他们的工作效率为依据来支付。

从心理学角度看，这种混合模式体现了许多调查执行机构得到的资助并不十分充裕的事实。资助不够充裕使他们无法奢望聘用那些调查速度较慢但调查质量很高的调查员。不论采用的付酬方式是什么，访谈的质量都必须是我们可以接受的，但混合式付酬可使我们认识到，在以标准方式付酬时，较之那些相对不太能干的调查员，许多能干的调查员不仅每小时完成的问卷数更多，而且拒访的数量也相对较少。据此，我认为若将酬金的多少和完成数量与拒访数量之比联系在一起，我们便可以建立起一种更为公平有效的回报机制。

在过去 10 年中，有必要时，NUSL 都会使用这种混合式付酬方式，在用这种方式付酬时，从未听到过有人抱怨根据他完成的调查业绩付酬是不公正的。当然，较之只采用标准的按小时计酬的付酬方式，混合式付酬方式要求自始至终都能对调查员的工作有所记录，但这种记录恰恰就是调查员个人能力和效率的反映，我相信它与我们对调查质量的期待是一致的。

总之，我们应通过培训使调查员明白，如果他们能够达到调查组织者对他们的要求，那么他们便会因此而得到更高的报酬。高质量的电话调查并非一件容易的事，正因为如此，一个优秀的调查员似乎也乐于使报酬在一定程度上反映自己的工作成果。那些对一般方法步骤和工作要求已经有所了解的调查员，不必每次都参加每个调查的第一阶段的培训。当然为慎重起见，参加一下也没有什么坏处。我认为即使是很有经验的调查员，虽然不必每次都参加，但有机会参加一次这样的培训，也可温故而知新。尽管一些有经验的调查员也许对此不以为然，但我还是想告诉这些人，偶尔回顾一下标准化访谈的方法和步骤是很有意义的，因为它能使调查员懂得，他们的经验水平的高低并不取决于宣读问卷的精确程度（参见 Groves，1989，pp. 383 ~ 385）。

最后还有一点：调查员不仅应该了解（且应该时刻牢记）他们有责任使所有他们所接触过的人都对调查留下好的印象。换言之，我认为调查员应从内心感到，他们对调查研究企业负有职业上的责任，因而必须努力避免使被调查人对电话调查留下不好的印

象。

为某一特定的调查项目进行的专门培训

不论调查员的经验有多么丰富,参加一般工作经验的培训有多么频繁,他们都必须参加为每一个新的调查项目举行的专门的培训。一般来讲,为这一目的而举行的第二阶段的培训应该按以下各个方面的内容来组织:

1. 阐述调查目的;
2. 介绍抽样库产生和电话号码处理的方法;
3. 介绍引导/选择单的使用法;
4. 介绍回撤陈述的使用方法;
5. 如果调查规定必须使用拒访报告表(RRF),应该对它的使用方法进行介绍;
6. 详细介绍问卷内容,包括它的实际使用方法。

在使用 CATI 系统进行调查时,不仅应清楚地说明任何特定软件的操作方法,而且应该给调查员安排足够的时间进行实际操作。

合乎规范的进行这一阶段培训的做法是,令每个将要参加实际调查的有关人员都能得到一套为每一个专门的调查准备的培训材料。这套材料应该包括要将调查中使用的表格的实例,作为在培训课堂上口头介绍内容的补充的书面指导。

调查目的

只有在了解调查的目的和意义之后,调查员才能认识到他们将要搜集的资料的重要性。正因为如此,我们认为调查组织者应该亲自参加这一阶段的培训,现身说法地介绍调查的目的。这样不仅仅是因为调查组织者最清楚调查的目的,更重要的是这样可以给调查员提供一个与他们最终为之工作的那些人交流的机会。

在对调查员解释调查目的时不必过于详细。此外,不要武断地决定是否应该与调查员分享研究假设(如果确有研究假设)。如果在某些调查中,诸如这样的知识有可能因调查员对调查数据有所企图而产生某种偏倚,为谨慎起见,我们还是不向调查员提供这样的信息为宜。但反之,如果调查员对调查目的毫不知情,则调查员的积极性也会因此而受到损伤。

抽样库和呼叫单

我们应该使调查员在学术研究的意义上对抽样问题有一定了解。这些问题包括研究总体的性质和大小，抽样库是怎样和在什么程度上代表了总体，以及调查使用的抽样方法是 RDD 还是以清单为依据等。在调查中使用的呼叫单必须在培训班上进行讨论。在讨论这一问题时，我们必须提醒调查员应该对回呼那些很难联系的被调查人的问题予以特别的关注，了解这些呼叫单上的电话号码，与其他他们曾经可能使用过的呼叫单上的电话号码之间存在的差别。调查员应该牢记，在每次拨号之前，他们都必须先填写日期、时间和他们的 ID 号码（在使用 CATI 系统时，软件会自动填写这些内容）。我们应对所有特定的结果处理码一一予以介绍。对其中调查员不熟悉的那些码，不仅必须介绍，而且还必须作详细介绍。

调查简介/被调查人选择单

此后，我们应该向调查员详细地介绍调查简介/被调查人选择单。依据选择方法复杂程度的不同，培训人员可以采用各种角色变换的方式对调查员进行这方面的培训。让一些调查员扮演可能的被调查人，而让另一些调查员扮演使用简介/选择单进行户内被调查人选择的角色（每次一个）。采用这样的培训方式将有助于调查员了解和掌握被调查人选择方法中的难点，当然前提是选择方法中的确存在某些难点。这样做的目的是使那些需要加强这方面训练的调查员，在进入下一阶段的培训或开始实地调查之前，有机会得到这方面的模拟训练。

回撤陈述

因为在调查员宣读介绍/选择单的时候，住户中在家的人和可能的被调查人也许会提出一些问题。这时常常需要调查员使用特定的回撤陈述作应答。因此我们必须训练调查员如何在这一时刻使用特定的回撤陈述。首先我们应该使调查员了解，各种标准化应答之间存在的细微差别。然后，我们还应该告诉调查员，他们必须在正式访谈开始之前，练习使用这些陈述。

拒访报告表

如果一个调查准备使用拒访报告表,那么我们便应该在介绍调查问卷之前介绍它。每次调查使用的拒访报告表,都会在一定程度上与调查员以往曾经使用过的那些拒访报告表有所不同,所以我们应该把介绍的重点放在这些差别上。与此同时,我们也应该告诉调查员,记录在 RRF 上的信息是很有价值的。如有必要,我们还应该告诉调查员如何在今后试图实现拒访转化时使用这些表格。

问卷

在此之后,培训员应该对调查员进行有关问卷内容和使用方法的培训。为了使所有的调查员能对问卷的内容和使用方法有具体的认识,培训者即使不能全部至少也应该逐项且完整地宣读问卷所列的大部分题项(包括答案选项)。跳答型题项即那些要求调查员根据被调查人选择的不同答案选择不同题项的题项,必须予以特别的关注。在用 PAPI 进行电话调查时,情况尤其如此,因为这时跳答题项的选择必须由人工确定。如果问卷中使用了开放式题项,那么培训员就应该告诉调查员,他们应该如何简明地记录被调查人给出的答案,包括记录的具体形式和字数。(在使用 PAPI 时,问卷上已经预留了空白。调查员可在上面记录被调查人给出的开放式问题的答案。预留的空间大小,实际上体现了我们对记录长度的期望。)对某些题项特有的那些注释或说明,培训员也应该一一向调查员解释清楚。

在详细地介绍了问卷之后,我们应该再一次用角色扮演的办法进行模拟的问卷使用练习。我们可将调查员分组,每组 4,5 个人,由督导员充当被调查人的角色练习问卷的使用。我们可用循环方式来练习跳答题的跳答选择,根据督导员给出的答案,一组调查员可轮流选择应该跳答的题项。(调查员一般都乐于参加以这种方式进行的这一部分内容的培训,因为这样他们都有机会主动参与。)一般来讲,角色扮演练习应该进行 20 ~ 60 分钟左右,具体的时间则应以问卷的长度和复杂程度而定。角色扮演的目的并不是要使调查员成为问卷设计的专家,而是要增加他们对使用的工

具的了解,进而消除他们在使用这一工具时可能产生的各种疑问。

结束培训

这种针对每一具体调查项目进行的培训,一般需要2~4小时。在结束培训之前,培训者应该简要回顾一下全部培训内容,并应使调查员明白,了解和掌握所有培训内容是十分重要的。培训者还应该间接地提一下,此后可能会进行的任何一种调查前或在岗培训的性质。在结束这一阶段的培训前,给调查员介绍一下整个调查的具体安排也是很有用的。

调查前实习和在岗培训

虽然培训的作用十分重要,可以使调查员了解调查过程中各个方面的情况,但最好的培训却是在实际调查过程中进行的,它的形式可能是调查前的实习或在岗培训。那些熟悉电话调查的人都了解,通常都需要经过一两个单元的访谈之后,调查员才能真正执行一项新的调查。基于这一认识,比较理想(但费用比较高)的培训策略是,令每一调查员在真正为最终的调查样本增砖添瓦之前,参加实际访谈的实习。(实习使用的电话号码应该是从抽样库中随机选取的一个子集,以免使最终的样本产生偏倚。)

如果一个调查研究者决定采用这种实习方法,那么他还会进一步发现,这将大大有助于他们发现(尤其是在一组新调查员中发现)那些可以搜集到高质量数据的优秀调查员。这样我们就可能在应聘的那些调查员完成实习(至少完成一份问卷)之后,才做出是否聘用他们的决定。例如,从某种意义上讲,错失一次计划中的实习,也许就是一种警示信号,它警示我们那个调查员可能具有某种潜在的不可靠性。所以在实习阶段,督导员的责任是不断地对调查员的各个方面工作的做出反馈(参见第6章)。反馈的依据是在他们完成每一份问卷之后立即进行的检查,对每张最终未能完成访谈的呼叫单进行的审查和花尽可能多的时间对实际进行的访谈的聆听。如果条件允许,调查员可以对访谈进行录音(如果被调查人同意的话),这样将有助于督导人员对录音重新演示的特定问

题做出具体评价(参见 Fowler & Mangione,1985)。因为这一工作对督导人员的要求很高,所以一个督导员可以照看的调查员最多不应该超过10个。在实习阶段和一个要求很高的调查的最初阶段,即使这样的比例可能还是过高,在这样的场合,调查员与督导员之比最好保持在4:1 或5:1 左右。

督导员和实地调查主管应当经常考察新调查员的能力,并据此决定是否继续在一个特定的调查中聘用他们。在实际调查过程中还应该对调查员的工作进行定期的考察。只要需要,督导员和实地调查主管都应该对那些不能满足调查规定要求的调查员进行个别指导,以使他们的工作有所改进。

对比较困难的调查,我们应当考虑使用结构性较高的个人化的信息反馈系统,以及时向调查员传达有关他们工作情况的各种信息。每月应该打印一次反馈信息。在仔细审读这些信息之后,我们可以了解哪些方面的问题已经得到改进,哪些方面的问题仍有待改进,这样可以使在岗培训更加有的放矢。

在理论上讲,在岗培训从来都没有停止过。监控调查员的访谈质量,进而对所有的调查员提供实时的信息反馈是每一个督导员永远不能推卸的责任。这一责任不会因为他们的资历不同而不同。在这样的培训由一些经验丰富和高度负责的督导员实施时,一批高度熟练的优秀调查员必将由此产生:那些调查员不仅对工作充满热情,而且严格遵守操作规范,他们决不会使调查数据产生偏倚而降低质量。

总调查误差和调查员

尽管每一调查机构都会动用一定的资源来聘用和训练调查员,但仍然无法完全避免与调查员有关的误差的产生(调查数据中与担当调查员的那些人自身某些特征有关的变差和/或偏倚),即使是质量最好的电话调查也不例外。因此我们奉劝从事电话调查的研究者,都应适当考虑那些会导致与调查员有关的误差的操作因素和统计调查问题。例如,一个特定的调查使用的调查员如果

越多，其与调查员有关的量度变差就可能越低，但与调查员有关的偏倚却可能有所增加（参见 Groves，1989，p. 315）。当然，在实际调查中，调查机构可以聘用的有酬调查员的人数必定会受到他们能承担的费用的限制，这一点就像我们在第 6 章中将要讨论的那样。

有几种可以估计问卷特定的题项中因调查员的影响而产生的变差的统计方法。凯思（Kish，1962）对如何用 rho（类间相关系数）来测量调查问题中与调查员有关的那些变差做了介绍（同时参见 Fowler & Mangione，1990；Groves，1989，p. 3 18）。沃尔特（Wolter，1985）则对"折刀"方差计算法的使用做了介绍。他认为在用这一方法计算调查统计量（如均值）的估计值时，每次都必须去除某一调查员调查的全部调查对象的数据。只有这样计算得来的统计量的方差，才可以用来估计数据中那些与调查员有关的影响的大小。水平较高的读者，可以读一读斯托克斯和明义（Stokes & Ming-Yih，1988）二人的有关著作。他们的著作对如何建立估计调查员对数据的影响的模型问题进行了专门介绍。此外读者也不妨读一读格罗夫斯（Groves 1989，pp. 360 ~ 381）的著作中论述这些问题的章节。

本书的所有章节中，我始终都在努力对降低总调查误差和测量总调查误差的数量这二者加以区别。这两个问题是高质量的调查工作中彼此相关的两个部分。就与调查员相关的误差而言，我们为减少它的存在可以做的工作是有限的。因此调查研究者应该试图比较精确地估计它存在的性质和数量，以更为确切地阐述他们的调查发现。

练　习

1）编制一幅某一电话调查机构招聘调查员的广告（备注：在广告中可包含一些图表）。

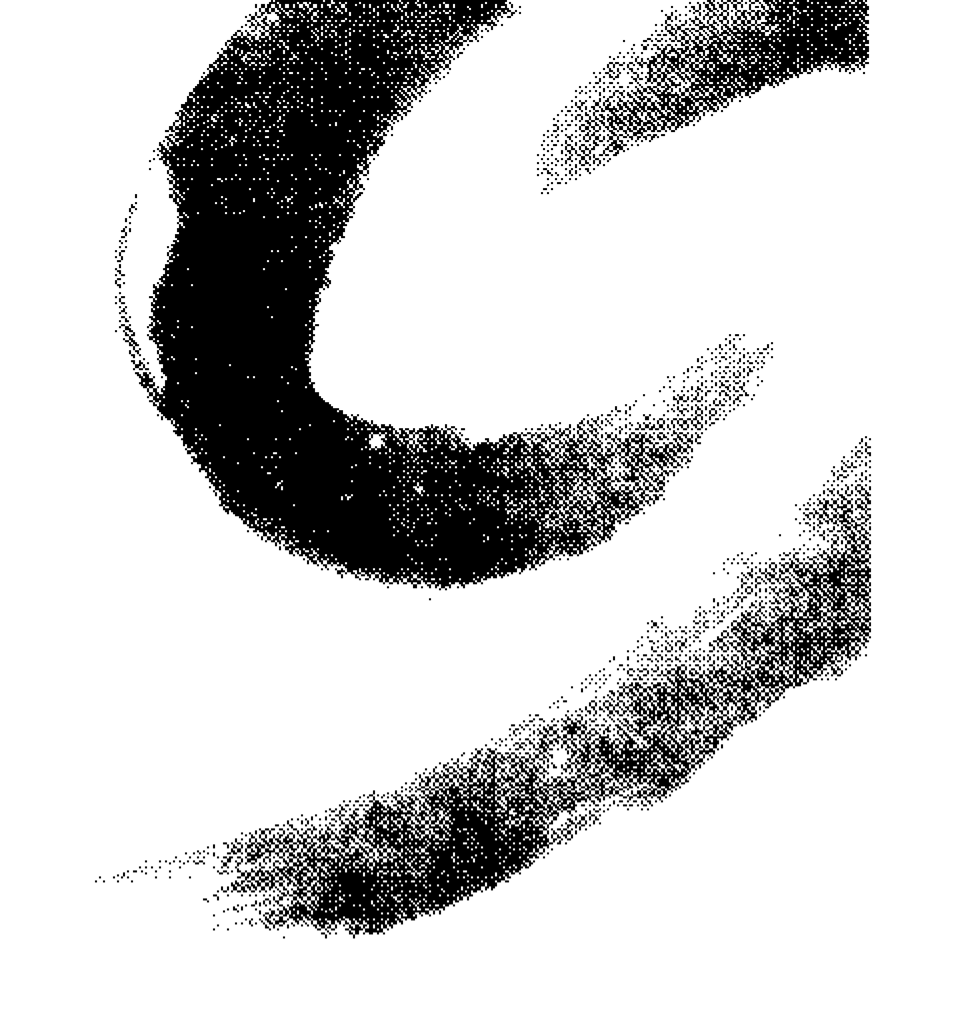
2）扼要地阐述使用有酬的调查员和无酬的志愿者进行电话调查有什么不同。

3）撰写一段用于在一个十分不愿意参加调查的被调查人和非

常善于说服(但十分有礼貌)的调查员之间进行的对话。

4)草拟一份3小时左右的,针对某一聘用20个新(有酬)的和20个有调查经验的调查员的特定调查的培训计划。试给每一部分的培训内容分配所需的时间(以分钟为单位)。

5)作为一项课堂作业,试将培训班上的所有成员编成若干小组,用角色扮演的方式来练习电话调查的问卷使用方法。



督导Ⅱ:安排督导的工作

Supervision II:Structuring Supervisory Work

本书尤其强调这样一点,那就是电话调查能通过对整个资料搜集过程进行监控给我们提供更多的降低总调查误差的机会。在第3章中,我们给读者介绍了一个高度常规化的处理抽样库中电话号码的系统是如何组织的,督导(特别是控制抽样库的督导人员)的重要性也在那一章中作了专门的介绍。第5章介绍了培训人员在培训工作中的作用,以及督导人员在在岗培训和访谈过程监控工作中的作用。

在本章除进一步讨论督导人员在整个数据搜集过程中的作用之外,我们还准备用一定篇幅来讨论一下在一个高质量的电话调查中,督导人员应该承担的其他各种督导工作,它包括实地调查主管的作用和其他各种保证电话调查顺利进行的行政管理工作等。

读者也许难以相信,但实际上却完全是可能的,有时我们只需要一个人便可担当一个电话调查的全部(至少是大多数)督导任务。例如,在将近10年时间内,我一直在教授新闻专业高年级本科生的民意调查课程,每门课历时2周,每年3次。作为这门课程的一部分,我安排学生作为调查员参加一个RDD电话调查(最终的样本量为400份左右的全部已经完成的问卷)。该调查采用户内被调查人系统选择法,要求对指定的难以联系的被调查人做最多不超过3次的重呼。调查使用的问卷平均包括20个左右的封闭性题项。我在10天之内做了以下这些工作:建立了一个抽样库,和学生一起编出了问卷,对学生按调查员的要求进行了培训,计划安排好每一阶段的访谈工作,承担所有访谈的督导工作,控制抽样库中电话号码的处理,录入并编辑全部调查数据,进行一组相当稳健的统计分析,并将分析结果给全班同学做汇报。

虽然本书的大多数读者可能不需要像我这样,集所有督导工作于一身,但有些读者需要担当的工作也许比这个还要多。例如,一个通过电话调查来搜集撰写毕业论文所需数据的毕业生,也许在上面各种工作之外,还必须独自完成所有的访谈。在另外一些场合,研究者可能委托某一调查机构进行电话调查,或可能考虑建立一个临时的组织机构来完成某一特定的调查。在这样的情况下,如果研究者熟悉电话调查的操作步骤和组织结构,无疑对集中组织管理电话调查和合理配置调查费用是十分有用的(参见Bass

& Tortora,1988;Berry & O'Rourke,1988;Frey,1989,pp.185~199;Groves,1989,pp.526~550)。

这时,我们需要考虑的是用于调查的操作系统是否要具有CATI这一性能。如果希望我们的操作系统具有CATI性能,那么调查需要增加的投资,就不仅仅限于购买计算机软、硬件。例如,完整地配备一套CATI设备,需要有比同水平的PAPI系统多33%~50%的办公面积。此外CATI还需要有更多的投资用于购配家具和办公用品,更重要的是它还需要有更多的预算用于聘用软件编程和硬件维护人员。尽管CATI在技术上的确很有吸引人的地方,但它并非包治百病的灵丹妙药,所以在装备这一系统时我们必须考虑周全,这一点我们已在本书第1章中给大家作过介绍。而迄今为止,我们还对采用CATI的电话调查肯定比采用PAPI的电话调查具有更高的成本效益这一问题知之甚少(参见Catlin & Ingram,1988)。(在这里,效益关系总调查误差的大小。)因此,我们认为对这一问题感兴趣的读者,应进一步阅读有关那些讨论在什么时候应该采用CATI,什么时候应该采用PAPI的文献。

不论操作结构是什么,也不论是否使用CATI,对所有的电话调查来讲,为了保证调查的质量,某些基本的督导工作都必须考虑。在这一章我们将要讨论的就是这样一些与督导有关的工作。

访谈阶段的人员和日程安排

为了安排好访谈阶段的工作,督导人员需要做的工作是很多的。实际上,如果这些工作没有得到很好的安排,那么访谈阶段的工作就不可能进行,如果进行了,也只不过是一种无组织的访谈。而在进入访谈阶段之前的诸多需要督导员完成的工作中,首先需要完成的工作是人员和日程安排。在这一节将要讨论的有关人员安排的大多数问题中,都是针对付酬的调查员而言的,不过其中某些问题同样也适用于无报酬的调查员。

人员安排

在大型调查机构中一般都设有实地主管一职,负责招聘和安排调查员。但在规模较小的调查机构中,这一工作通常则由某些负有督导责任的人员兼任。为了确定调查的实地访谈阶段究竟需要配备多少名调查员,我们必须考虑以下几个方面的问题:

1. 每名调查员在每一工作单元可以完成的问卷数;
2. 每名调查员每周可以参加多少个单元的访谈;
3. 每一单元能参加调查的调查员的平均人数;
4. 每星期将安排多少个单元的访谈;
5. 计划完成的问卷总数。

例如,如果我们使用 RDD 抽样,完成一份问卷需要的时间在15 分钟左右,那么一个调查员一小时可完成两份问卷。如果每工作单元为4 小时,那么一个调查员在一个工作单元共可完成8 份问卷。假如我们每星期安排6 个工作单元,共有电话10 部,那么每一星期总共可完成问卷480 份。根据调查员每周平均参加工作的次数,我们就可估算出使电话设备能保持全员工作状态所需的调查员人数。调查要求完成的问卷数越多,对这一估计数所做的修正(增加)的可能就越大,因为在这种情况下,出席率相对较低,而减员或其他消耗则可能较多。例如在这一例子中,调查共需完成的问卷数为2 400 份,在全员工作的情况下,访谈需延续 5 个星期左右。

如果调查员平均每周完成3 个工作单元,那么在5 个星期的访谈中共可完成120 份左右的问卷。[即:8(将每一单元完成数)×3(每周工作单元数)×5(访谈阶段持续的星期数)]。严格地讲,这一调查只需调查员 20 名,便可完成2 400 份问卷。但为保险起见,考虑到减员和出勤率有可能低于预期的水平,我们应该聘用和培训的调查员人数,应比这一数字高10% ~30%。

诚如前述,如果一个调查聘用的调查员人数相对较多,那么从理论上讲,这个调查就有可能在一定程度上使与调查员有关的方差有所降低(Groves,1989)。但是如果我们在这一问题上采取了极端的态度,不适当地过多地增加调查员的人数,就有可能使我们试

图为降低总调查误差而投入的成本与产出的效益不成比例。格罗夫斯(Groves,pp.62～72,365～374)从成本效益的角度对适度的调查员(和督导员)人数问题作过比较深入的讨论。虽然在这一领域内,我们现在所掌握的知识还是不精确的,还难以据此对一个调查项目需要的调查员的确切人数做出精确的估计;但这些知识已足以使一个实地主管在考虑到总调查误差并兼顾实际需要和成本二者的前提下,对项目实际可能需要的调查员人数做出大致的估计。

为了能招聘到能干的调查员,我们应该张贴一些招聘广告,而最适合张贴招聘广告的地方是高等院校的校园或校园附近的某些地方。在这些地方张贴招聘广告,无疑会吸引学生、学生和教职员工的配偶及这些机构所在的社区的居民前来应聘。虽然学生一般都可以成为优秀的调查员,但是聘用他们作调查员通常也会有一些不足之处。因为他们都承担一定的学业,加之学生特有的生活方式,因而出勤率不太稳定。所以我始终认为,如果我们想使调查的实地访谈阶段的成本效益达到最高,就不应当只考虑从学生中聘用调查员。

正如我在第5章谈到的那样,决定是否聘用一个调查员的最重要的因素是调查员在电话上的表现。因此考察者对申请人最初进行的考察都应该通过电话进行。申请人与高质量访谈可能有关的那些语言习惯,如讲话的速度和声调等都应该在考察之列。考察的内容应该包括讲话是否有自信、语调是否悦人、吐字是否清楚、讲话的速度是否适当、讲话是否机智和是否显得比较成熟等等。为了可以给申请人在这方面的能力分出等级,申请表中应该包括一些开放式问题,使应聘者能有回答问题和谈话的机会。

工作日程安排

在人员安排已经初步完成,并对他们进行了培训之后,我们必须对工作日程做最后的安排。对一般人口总体的调查,虽然各个社区的居民有着各自不同的作息方式,在家或不在家的时间也不尽相同,但最有可能与大多数可能的被调查人取得联系的时间是星期天到星期四的晚上和星期六下午(参见 Frey,1989,pp.232～235)。例如,在美国南部地区,许多居民有在星期天晚上去教堂聚

会的习惯。如果计划每单元安排 4 个小时的访谈,那么我们建议将晚上那一单元的访谈安排在下午 5:00—9:00 或 5:30—9:30 之间进行(参见 Weeks, Jones, Folsom, & Benrud, 1980; Weeks, Kulka, & Pierson, 1987)。即使如此,它还是与晚饭时间重叠,但调查员应该可以利用这一机会,在被调查人比较方便的时候安排一次重呼,一个有经验的调查员做到这一点应该是不成问题的。如果一个调查的抽样区域跨不同的时区,那么访谈时间的安排也应该进行相应的调整。

星期天到星期四晚上这一段时间,工作效率通常比较稳定,但也有些时候例外。例如秋天正是美式足球赛季,一部分公众不希望他们在电视上观看职业足球比赛时被人打扰。星期五和星期六晚上的工作效率最低,因为大多数社交活动都在那两个晚上进行。不仅如此,调查员和督导员本人也不太乐意在那两个晚上工作。

在大多数地区,访谈效率最高的时间是星期六上午 11:00 到下午 4:00。尽管在这一时间打电话联系,有时可能需要进行一些重呼,因为时值周末,指定的被调查人常常会有一些家务或杂务需要在那一时间处理。经验告诉我们,那些在工作日的晚上联系不上的可能的被调查人,有可能在星期天联系上。不过在一年之中某些时间的星期六和星期天下午,访谈效率可能会有所下降,因为那时电视正在转播体育比赛。

还有些时候,大部分公众都在观看特别的电视节目(如超级保龄球大赛),这时通常也不应该安排访谈。这些时候不仅访谈效率比较低,而且还会出现比平时更多的拒访。此外,安排在这些时候做访谈,访谈质量也会有所下降,测量误差则会因此而增加。

根据实地调查阶段允许延续时间的长短,如有必要和可能,也可以在工作日的白天安排一些访谈单元。一般来讲,白天进行的访谈,应该安排在下午 1:00—5:00,而不要安排在上午。在工作日白天进行的访谈,虽然有可能因为指定的被调查人在单位工作而无法与他们取得联系,但它有助于我们清理 RDD 抽样库中的办公或非居民户的电话号码。经验告诉我们,RDD 调查中,工作日下午调查员的工作效率(每小时完成的问卷数)大约是晚上的 60% ~ 80%。把访谈安排在工作日的白天的一个优点是,年长的被调查

人在那时接到来自不相识的人的电话似乎不那么不安，虽然比较警觉，但比较配合。此外，没有外出工作的父母，因为身边没有学龄儿童和/或配偶而感到更为放松。如果只安排白天的工作单元，那么便会因为无回答（无法联系）误差，而无法得到一个合乎规范的成人样本。

目标总体非一般公众的电话调查，访谈应该安排在比较容易与被调查人取得联系的时间进行。例如，在工作地点进行的律师或其他职业总体的电话调查，应该将访谈集中安排在工作日的白天，并允许进行多次重呼。而在家中进行的高中学生的电话调查，访谈时间最好集中安排在星期一到星期四下午 3:00—7:00。如果样本抽取的总体在可以接受访谈的时间这一问题上，的确存在着某种特殊性，我们在安排访谈日程时就必须对这一问题加以考虑。

在安排访谈日程时，季节性变化也是我们应当考虑的问题。如果我们有理由确信，调查不宜集中在一年的某一时间进行，那么我们在日程安排上就应该把这一问题考虑进去。对于那些实地调查时间相对比较短的一般人口总体的调查，最好不要把搜集数据的时间安排在夏天或十二月中下旬（参见 Frey，1989，p. 234）。否则无回答率就可能上升，因为那时许多人都在度假，或安排了必须在那一期间处理的各种私人事务（如为节日准备食品或购物等）。

一旦我们已经对访谈时间做出了具体的安排，就应该立刻把人数足够的调查员落实到每一部用于调查的电话上。如果在访谈进行了一个星期之后，我们发现至少有一部电话，因调查员未能按时参加调查（这种情况时有发生，因为电话调查的调查员大多是临时工），而未能经常投入使用，那么我们考虑采用“超额订购”调查员的办法来解决这样的问题。例如，如果可用于调查的电话为 10 部，那么我们可在每一调查单元计划安排 11 名调查员，其中有一名作为机动。经验告诉我们，在大多数情况下，至少总是有 1 名调查员不能按时参加调查。当然如果所有计划安排的调查员全部都来参加调查，那么就有可能会产生一些问题。

当然不同的实地调查需要的调查员人数不同，如果需要人数很多，计划安排调查员的工作可能会是一个非常费时间的工作。我们建议由专人（最好由一名有一定督导经验的人）负责与计划安

排作实地访谈的调查员进行联系,以了解他们是否可以参加实际上已经做了安排的访谈单元的访谈。如果某一位调查员一直无法保证正常参加已经安排好的调查单元的工作,那么我们可以通过电话给他/她一个警告,告知他/她的聘用有可能终止,以便我们能更好地计划安排访谈单元的工作。

在计划安排调查日程时,最后一个可能要考虑的问题是,在某些情况下我们在安排访谈日程时,需要考虑调查员的某些人口学特征。如某些调查单元只能安排一组有某种人口学特征(例如女性)的调查员。诚如前述,有时调查员的人口学特征可能与调查的主题(例如先前提到的例子中的性骚扰问题)有关,因此在安排调查日程时,我们必须考虑使调查员的人口学特征与调查的需要匹配(参见 Groves,1989,pp. 398 ~404)。如果我们因为难于做到这一点,而在计划安排访谈日程时没有考虑这一问题,那么就相当有可能导致测量误差的增加。只要我们能使调查员明白,我们之所以只安排有某些人口学特征的调查员进行某些调查,并非是对其他调查员的歧视,而是因为调查的需要,那么因此而可能在调查人员中引起的问题便可以减少到最少。

访谈单元的组织安排

除了系统高效地安置调查员和组织安排调查日程之外,督导人员也需要逐天组织安排集中化的电话调查的各项工作,如调查员进行电话访谈的地点等。这一工作一般应在每一访谈单元开始之前,且假定在上一阶段的工作结束之后的某些工作已经完成的前提下进行。尽管督导人员在这一阶段所需做的工作,在 PAPI 和 CATI 两种不同的调查环境之间存在某些差异,但也存在许多相似之处。

在调查员到达工作地点时,他们首先应该很容易就找到为他们安排的进行电话访谈的房间。他们的工作站应该光线充足、整洁和井井有条。如果是 PAPI 调查,那么呼叫单、问卷和铅笔等调查必备的物品已经一应俱全。最理想的是每部电话和/或 CATI 工

作站都安装在一个单独隔断或小单间中,以使每一调查员都能有一个属于自己的相对封闭的工作环境。如果我们无法提供这样的工作环境,那么访谈设备可以安装在一张张桌子上,但这些桌子的安放形式,应注意不要使调查员相视而坐。

在动手准备 PAPI 访谈单元的房间之前,我们必须将前一单元的呼叫单,以最后的处理结果码为依据进行分类整理(人工进行)。而那些仍然还存活的号码(如某些需特别进行回呼的号码和通了但没有人接听的号码)将被插进为数不多的尚未呼叫过的呼叫单中,以备在下一单元调查时使用。而在全 CATI 系统的电话调查中,我们一般可以用软件按照调查的需要将电话号码分配给调查员,且令计算机根据每一号码的呼叫史,在比较恰当的时间尝试进行重呼(参见 Groves,1989,pp. 196 ~201)。

在大多数调查中,不论是 PAPI 还是 CATI,电话号码一般都应当以某种随机的方式分配给各个调查员。这样调查员就可以工作在一个活动的抽样库的某些子集上。而"随机化将加强公正性"(Groves,1989,p. 362),调查员的工作热情也会因此而增强。有时,督导人员需要将某些特定的呼叫单分配给某些特定的调查员。这时督导人员应该使调查员了解,他们之所以要这样做的基本原因是,即使使用将抽样库中的电话号码自动分配给调查员的全 CATI 系统,为了能将某些特别的电话号码分配给合适的调查员,也难免需要进行一些必要的人工干预。

在做 PAPI 调查时,最好不要在一开始就给调查员分配太多的呼叫单,而是应该由他们自己根据需要再从督导人员那里得到更多的呼叫单。因为从某种意义上讲,调查员对新的呼叫单所提出的每一次要求,都意味着他们在呼号过程中取得的一个进步。我们之所以要求由督导员来控制 PAPI 调查的呼叫单的流向的另一个原因是,如果我们允许由调查员自己来选择呼叫单,那么呼叫单的选择便会受到调查员自身行为方式的影响,从而可能导致抽样偏倚和/或增加与调查员有关的误差。

经验告诉我们,负责安排 PAPI 访谈阶段工作的人都必须了解这样一个问题,那就是一个中等水平的调查员,在工作单元的前半段究竟可以处理多少张呼叫单。显然,这一问题主要取决于问卷

的长度和抽样库的效率。如果问卷较长和/或抽样库效率比较高，那么调查员势必要在访谈上花费很多时间，而用于拨号的时间就比较少，这样每小时处理的呼叫单就相对较少。相反，如果问卷较短和/或抽样库效率较低，那么就需要用更多的时间来清除不合格的电话号码和无法接通的样本，因此每一小时需要处理的呼叫单就比较多。在 PAPI 和 CATI 调查中，如果我们使用的问卷比较长（在 20 分钟，或 20 分钟以上），督导员应该告诉调查员，在访谈单元结束之前，什么时候才可以不再进行下一份问卷的调查。如果督导员没有明确地提出这方面的要求，在进行较长的问卷的访谈时，调查员常常会在访谈工作还应该继续进行的时候停止工作，从而增加了总调查的费用（参见 Groves，1989，p. 491）。

在 PAPI 调查中，一种行之有效可以提高工作效率的办法是，在每一工作站摆出的问卷数少于整个工作站调查员可以完成的问卷的平均数。我们可以用这种方法促使调查员更好地工作。因为站中的调查员将会在工作单元结束前，看到他们已经完成了放在工作站中的所有问卷，必须从督导员那里再得到几份问卷。无疑这会使他们有一种成就感，进而对他们的工作有所促进。

依据我们使用的电话设备的型号的不同，有些清理工作必须在工作单元开始之前完成。至少，在调查员来到之前我们应该把每一个工作站都打扫干净。如果用于调查的电话都配备了单独的耳麦，那么督导员就应该告诉调查员，他们应该自行保持耳麦的清洁。大多数型号的耳麦上的麦克风都不会与调查员的嘴发生接触，因此只需要偶尔对它们作一下清洁就可以了。如果我们使用的是手提电话接听器，那么接听器上的整个外表面都应该在每一工作单元开始之前，用无毒的溶剂加以清洁。此外，每星期还应对它至少做一次彻底的清洁，即将每一个接听器都拆卸下来，不仅清洗表面，而且也清洗内部。

保持工作环境的清洁固然是对调查员的礼貌和尊重，但更重要的是整齐清洁的工作环境有利于调查员的健康。调查员的嗓音对访谈工作的成败至关重要，感冒和其他呼吸道疾病都会影响他们的嗓音，使他们不能很好地完成访谈工作。如果说无法保证稳定的出勤率本来就是许多临时聘用的调查员身上经常发生的问

题,那么督导人员就不应该再使他们因健康问题而使这一问题变得更加严重。

在每一工作单元开始之前最后一个需要准备好的工作是,在调查员的文件夹中准备好针对每一个调查员的有关信息。在调查员进行电话访谈的地方附近应该安放一个文件柜或一组邮件箱,每一个调查员都应该在里面有一个单独的文件夹。调查员只要向里面扫上一眼便可以找到自己的文件夹,并在拨电话之前从里面发现有没有新的信息。用这一办法,督导员可以将一部分问卷分配给特定的调查员,要求他们进行重新处理。发给每一个调查员的备忘录,也可以像从一个调查员以前的工作中发现的个别问题而写的注意事项那样,放入每一个调查员的文件夹中。如果有些问题比较严重,需要督导员在下一阶段的调查开始前与调查员面谈,那么我们也可在文件夹中放上一个通知,叫调查员在开始下一阶段的工作前来面见督导员。而在使用 CATI 系统时,我们可以使用"电子文件夹",这样调查员便可以在开始电话访谈前登录到那里。

如果在调查员到达一个已经高度集中化了的工作地点时,那里的一切在他们看来(且的确)是井井有条的,这将肯定有助于使他们认识到,他们之所以来到这里,不是为了其他什么事情,而是从事一项具有专业水准的高质量的访谈工作。督导员或其他管理人员,作为调查员的另一方,他们的责任不仅是把工作单元安排得更有条理,而且还有责任把调查员从闲杂事务中解放出来,使他们能集中精力从事高质量的访谈,而不是"家务劳动"。

PAPI 和 CATI 调查中访谈阶段的督导

控制和提高电话调查的质量有做不完的工作,督导员在访谈阶段的责任更加重大。在这一阶段,他们的责任是保证抽样的完整性和搜集的数据的高质量。正因为如此,我们必须聘用那些精力充沛和经验丰富的人来担当这一任务。他们的报酬也该相应高一些,一般应为最低工资的 3 倍,1992 年大约为每小时 13 美元。

那些接受过社会科学研究生程度训练的人通常都了解,一个优秀的督导员应该具备什么样的素质。他们必须比较聪明,乐于在高度结构化的环境中工作,并且通常都十分了解科学研究的意义,以及高质量的抽样和访谈对降低总调查误差的作用。

一般来讲,在既考虑费用又考虑数据质量的前提下,比较恰当的督导员和调查员比率是 1 个督导员负责 8 ~ 10 个有经验的调查员(参见 Groves,1989,pp. 61 ~ 62)。如果一个督导员负责的调查员人数超过 10 个,时间就会太紧张,难以保证督导的质量。然而无论是出于成本效益的考虑,抑或实际工作的需要,都没有必要使一个督导员负责的调查员人数低于 8 个。惟一的例外是,在调查员都是新手且在实际调查阶段刚开始时,为谨慎起见最好将督导员和调查员的比率保持在 1:5 左右。

实地调查主管的职责是对督导员进行督导。这一职责包括督导员的聘用、培训和对他们的实际工作情况经常进行检查。督导员应该知道,调查主管可能会在事先不作任何通知的情况下,到调查现场检查他们是否很好地担当了他们应该担当的职责。督导员的职责因调查而异,下面我们将对这些职责作具体的介绍。

创造高质量的访谈环境

督导员的责任是保证由他们担当督导工作的那些调查单元的访谈质量。调查员应当清楚地认识到督导员的确负有并正在担当起这样的责任。我们建议督导员应该比调查员至少早 30 分到达工作地点,以便利用这一时间检查开始一个工作单元所需的准备工作是否已经就绪(如果尚未就绪,立刻使它就绪),并使自己做好一切准备,等待调查员的到来。

督导员在自己负责的工作单元应充分展现自己的权威。这种展现应该是自信、公正和首尾一致的。督导员的行为举止不应该令人感到紧张,但又不应过分友善。督导员应该以调查员更有可能从心里乐于接受的方式,对调查员的工作提出建设性的意见。如果督导员以一种过分草率、全盘否定或令人难堪的方式表达自己的意见,即使这种意见是建设性的,也可能为调查员所拒绝。

督导员应加强自己的能力,使自己能更好地解决在调查过程

中发生的各种与调查员有关的问题(Cannell & Oksenberg,1988)。这些问题也许是调查员虽然受到了足够时间的培训,但对调查技巧仍然不十分熟悉;也许是调查员虽然知道应该做什么,但不知道应该如何去做;或者也许是调查员虽然知道他要做什么,但没有足够的能力去做。这时,除非督导员能确切地知道问题的症结所在,否则她/他无法就遇到的问题给调查员提供适当的解决方法。

验证问卷完成情况

正如第5章所述,在PAPI调查中,调查员在完成每份问卷之后,必须立即对它进行快速而细致的检查,并在检查完后立即将它送交督导员。督导员在收到问卷之后,应该对整个问卷进行逐项检查,验证问卷中是否有问题遗漏,答案是否已经全部记录在案。那些记录开放式问题的答案的文字是否简单明了,易于阅读。为了能准确而迅速地完成这一工作,督导员必须非常熟悉调查使用的问卷。

在CATI调查中,验证问卷的工作由软件代劳,而不需要督导员在问卷上逐一验证。这里必须注意的一个问题是,有时督导员会觉得他们在PAPI中得到的监控工作单元进程的经验,在CATI环境中似乎"失去了感觉"(参见Berry & O'Rourke,1988,p.473)。(这是一个有待进一步研究的领域,通过进一步研究我们才能确定CATI和PAPI的优点和缺点究竟是什么。)

除非督导员已经解决或已经回答了调查员提出的问题,否则在PAPI调查中,验证已经完成的问卷是督导员的首要工作。验证的目的在于能尽早发现访谈过程中发生的问题,并在问题发生之后立即与调查员商讨补救办法,以免同样的问题再次发生。在许多情况下,并不需要调查员再次与被调查人进行联系,发现的问题就可以得到解决。然后,在某些情况下,关键的资料完全丢失,这时督导员必须督促调查员立即再次与被调查人联系,尽快将那些丢失资料补回来。

在PAPI调查中,问卷验证合格后,督导员应在附于问卷上的呼叫单上签字,以便他人确认验证人。实地主管应该不定期地对督导员验证过的问卷再次进行验证,以检查督导员的工作质量和

工作态度。不仅如此，在PAPI调查中，在完成问卷验证之后，但在数据输入计算机之前，我们还要对问卷进行编辑/编码。在这一过程中，一旦发现了问题，我们便可根据呼叫单上的签名，找到那个负责问卷验证的督导员，了解并解决发现的问题。

诚如我们在前面已经指出的那样，督导员和调查员之间的关系过于亲密势必会影响调查数据的质量。督导员必须经常对调查员的问题给出口头或书面的反馈。在调查开始阶段，在岗培训尤其重要。督导员给调查员的反馈必须是及时的和直截了当的，但分寸必须得当。督导员应该懂得，他们永远也不应该使一个调查员在其他调查员面前感到难堪。如果督导员采用递条子的方式来纠正调查员的小问题，或约请调查员在下一个访谈开始前面谈，就可避免这样的情况的发生。（在有些CATI系统中，督导员可以与调查员进行在线交流。）总之，督导员必须有权威和判断力，这样他才能在调查员已经发生问题但又不能或不愿加以纠正时，停止他继续做访谈单元中的其余访谈。

虽然要求调查员和督导员不犯任何错误的想法是不现实的，但是假如一个督导员在验证PAPI问卷时精力能高度集中，那么发生的错漏问题可能就不会太多。同样，如果调查员意识到他们做完的每一份问卷都将受到严格的检查，就会在工作中加倍小心，以认真的态度做好每一次访谈。

在采用拒访报告表的PAPI调查中，我们应该建立一个与其他检查系统类似的系统，对填好的拒访报告表进行验证。调查员应将填好的拒访报告表连同呼叫单一起立即递交给督导员。督导员在收到之后，应立即对填写的情况进行检查，如发现问题，应当立即反馈给当事的调查员。

对访谈过程进行实时监控

只要可能，电话调查就应该使用集中化的电话装备，以便督导员能在自己的电话上监听所有调查员使用的电话线路。有这样一种特别的电话，使督导员既可以对正在进行中的访谈进行监听，又能不为调查员和被调查人所察觉（但在1992年，美国参议院曾有一项立法提案，建议将这样的行为定为非法）。如果督导员使用的

是一部与调查员使用的每一条线路相连的普通电话机,他也同样可以进行这样的监听。在后一种情况下,督导员应在进入某条线路之前摘下话筒,否则的话,调查员(一般都不是被调查人)就会意识到督导员正在监听他们的电话。用这种方法进行监听,可能会有一些不足(线路的音量会稍微有一些下降),但它也有一个优点,那就是调查员可以由此知道他们的访谈正处在督导员的监控之中。

在 CATI 调查中,对正在进行中的访谈进行实时监控是督导员的主要责任,因为他们不需要像在 PAPI 中那样,验证已经完成的问卷。在 PAPI 中,实时监控虽然居于次要地位,但它仍然是督导员的职责之一。无论 CATI 抑或 PAPI,监控都可分为正式和非正式两种,前者采用结构化的调查员监控表(IMF),而后者则不采用任何结构化的方法。

督导员不必监听全部的访谈,而是应当系统合理地安排对访谈的监听。一般每次监听的时间只需几分钟,监听的对象虽然应该包括所有的调查员,但重点应当是那些经验比较少的调查员。对他们的监听不但应该更为经常,而且每次监听的时间也应该更长一些。我们甚至可以采用概率抽样的办法选择在什么时候、对什么人进行多长时间的监听(参见 Couper,Holland,& Groves,1992)。

调查员和被调查人之间进行交流的所有方面都在监听之列。这些方面包括:调查员使用的简介、被调查人的选择顺序、回撤陈述和问卷本身。根据调查可资利用的资源情况不同,我们可为一个调查"量体裁衣"地设计一个供督导员使用的 IMF。一张 IMF 可以:(a)帮助督导员给调查员提供实时的回撤陈述文本;(b)给实地主管提供有关调查员实际工作情况的资料;(c)给督导员提供某种形式的数据,这些数据对减小在调查题项中与调查员有关的测量误差有很高价值。

督导员应对这两个问题予以特别的关注。首先,在遇到被调查人给出的答案不够完整、模糊不清或与主题不相干时,调查员是怎样进行探询的。其次,在被调查人遇到可能对问题中的某些词语的含义不太理解时,调查员能否以无偏的方式准确地再次宣读问题,并定义/澄清这些词语。对调查员任何可能影响被调查人对

问题答案选择的所作所为(口头上的),督导员都应该予以密切注意,因为这样做也许会导致答案的偏倚。例如,无论在何时,即使在被调查人提出诸如"你是怎么考虑的"这样的问题时,调查员也不应该表达自己的观点,以免对被调查人产生不好的影响。调查员应时刻保持警觉,不要养成说类似"这太糟糕了"或"听到您这样说,我真是太高兴了"这样的话的习惯。

某些 CATI 系统可以对实际的访谈过程,包括调查员与计算机之间的人机互动过程进行实时监控。这样督导员就可以对访谈过程中调查员能否正确记录封闭式和开放式题项的答案这样的细节进行评估。而大多数 CATI 系统,虽然都不允许调查员输入的答案超出规定的范围,但无法识别那些没有超出规定范围可是不正确的答案。

大多数督导员都喜欢使用 IMF,即使是比较简单的也可以。因为它将有助于他们在调查单元中组织安排好自己的工作。在条件不允许他们使用 IMF 时,督导员应培养自己具有采用与 IMF 同样标准的对访谈进行非正式的实时监控的能力,并对调查员的访谈工作做出正确的实时判断。

在使用的电话调查系统不允许对调查员的访谈进行的实时监听时,督导员应该经常到正在进行电话访谈的房间走一走,照例在每个调查员前停留一会儿,简单听一听调查员正在进行的访谈。任何有电话调查经验的人都知道,在由一个有经验的督导员负责对访谈过程进行监察时,调查员再想对访谈做假是一件很困难的事。督导员应该有能力根据应答出现的频数和一致性/不一致性判断一个访谈的真伪。虽然允许督导员对访谈进行实时监控的电话调查系统更可取,但不论这样的监控设备是否存在,对访谈进行实时监控都应该是督导员的职责中不可或缺的一部分。

在访谈实际进行之初,对访谈过程的实时监听应该更为普遍一些。这时,每 8 ~ 10 个调查员最好能配备 2 名督导员。在访谈阶段的后期,监听工作也不应停止,即使是对那些有经验的调查员也必须如此。除非因为一个既定日程的需要,监听工作可以暂停,若没有特殊的情况,无论 PAPI 调查还是 CATI 调查,督导员都必须根据以往的经验分配一定的时间对访谈进行实时监听。

对呼叫单处理结果进行监控

在访谈进行过程中,对抽样库的处理进行监控也是督导员应负的责任。督导员应该随时对调查员使用完全拒绝和部分拒绝这两个结果处理码的情况进行检查。此外,他们还应该检查调查员对其他处理结果码的使用情况。同样,在做这些检查时,督导员应把注意力主要集中在新来的调查员身上。在 PAPI 调查中,督导员将通过人工整理检查呼叫单来实现对结果处理码使用的监控。而在某些全 CATI 系统调查中,一般不用人工而是通过根据督导员和实地主管的要求生成的各种不同形式的电子报表,对调查员处理抽样库的进度进行监控(参见 Weeks,1988,pp.418~419)。

有关抽样库处理还应该考虑的一个问题是,如何在访谈阶段重复使用某些电话号码。因为在实际访谈过程中,前一个访谈单元可能会留下一些需要在特定的时间和/或需要有特定的调查员重拨的电话号码。在某些 CATI 系统中,这一任务中的大多数或全部均可由软件完成。而在 PAPI 调查中,则必须由督导员亲自将这些呼叫单在合适的时间交到合适的调查员手中。

有时,有些调查员对继续完成前面的访谈单元遗留的那些没有全部完成的问卷缺乏自信,不愿再次与这些被调查人联系。这时,督导员应和这些调查员进行交谈,搞清问题的症结所在,然后再决定是否安排其他调查员完成这些还没有全部完成的问卷。

解决问题和回答问题

在整个访谈阶段,督导员必须不断地解决和解答可能遇到的各种问题。他们必须知道如何回答调查员提出的问题。他们不仅要能解决事先能预见到的问题,还必须随时准备解决各种意料之外的问题。有时他们还要缓解调查员之间存在的某种紧张关系。而在另外一些时候,调查员可能会过于饶舌,在访谈中讲过多的废话,从而影响了工作效率。督导员偶尔还需要与不安或发怒的被调查人直接交谈,为调查员(可能)的言语不当而向他们道歉,或向他们解释一些调查员没有解释清楚的问题。有时督导员甚至必须十分有礼貌地聆听那些希望和他们谈话的被调查人的谈话。

接受调查委托人的考察

调查委托人完全有理由到资料搜集的现场,对他们需要的资料的实际搜集情况进行实地考察。因为正是这些调查委托人,支付了调查所需的一切费用,所以无论他们想得到有关调查的哪一方面的第一手材料都是正当的。一般地,他们希望了解的情况是:调查员如何使用问卷和其他调查资料进行访谈,以及被调查人如何回答问卷中的问题等。但问题在于,调查委托人有时对被调查人在参与一项调查时应该表现出来的热情,常常抱有不切实际的期望。那些天真的委托人常常会想当然地认为,所有的被调查人都会对问卷感兴趣,且都会认为问卷上的问题对他们是很重要的,于是他们便会自然而然地假定,如果某些被调查人对调查不专注和/或不感兴趣,便一定都是调查员的过错。

在遇到这样或那样的与调查委托人检查访谈现场有关的问题时,督导员固然必须随时听取他们的意见,但在实地主管改变有关规定之前,督导员不能因委托人的意见而背离调查程序规定的标准。在任何时候,督导员都不应该在现场与委托人争论,应该让他们就他们所关心的问题直接与现场主管交谈。在此之后,如有必要,现场主管将授权督导员进行他们所授权的任何更改。

复查完成的访谈

还有一件督导员必须完成的工作,那就是对已经完成的访谈进行复查。在每一访谈单元,督导员除必须在访谈过程中对当时完成的每一个访谈进行检查核实之外,还应该在一个访谈单元结束时,通过回呼对部分已经完成的访谈进行复查。任何一个电话调查在编制预算时,都应该把这一笔费用包含在内。(在 1990 年,受某一律师事务所的委托,NUSL 进行了一项电话调查。该事务所希望将调查结果申报作为联邦法院审判的证据,因此在调查结束后,该事务所聘请了另外一个调查机构,对 NUSL 完成的全部 1 600 份问卷进行了复查。)通常复查都由督导级的调查管理人员负责实

施。他们必须通过电话与被调查人取得联系。首先,必须核实的问题是是否确实进行过一次访谈,然后再对问卷上记录的某一部分题项的答案进行核实。

如果我们能按照这一章提出的那些要求,在实际调查过程中对访谈进行严格的监控,那么在访谈单元结束时,可视调查进行的实际情况决定是否再对问卷进行效验。根据我们的经验,对新手所做的访谈,特别是他们在实际访谈阶段的最初几个单元所作的访谈,即使我们在调查过程已经做了严格的监控,在调查结束时,最好也能进行效验。此外,经验也告诉我们,最好能让调查员知道,他们完成的问卷都必须接受效验,但不必让他们知道有多少份问卷要被效验。

任何一份没有进行过复查的问卷都可能留有令人担忧的隐患。虽然它们并不一定意味着调查员的工作一定有什么不合适的地方(如有时,有些被调查人会莫名其妙地否认他们曾经参加过调查),但他们毕竟在提醒我们,我们还需对某一调查员的工作做进一步的检查。在 PAPI 调查中,有时一些粗心大意的调查员,可能会将呼叫单错误地附在另一份已完成的问卷上。这样,督导员在效验问卷时,就会拨错电话号码。

在采用 IMF 生成调查数据时,研究人员可以用来自效验表的数据对调查的测量误差进行研究,并确定它是否与调查员或被调查人有关。

结　论

诚如福勒(Fowler,1984)所言,专业调查研究人员对在数据搜集过程中与人有关的那部分误差的关注还远远不够。但通过这一章,以及全书其他章节的介绍,我想读者应该已经了解,电话调查的质量包括抽样的质量和实际搜集到的数据的质量,都与调查实施过程中督导质量有关。

许多调查机构,特别是商业性的调查机构似乎不太愿意建立一种强调和奖励高质量的电话访谈体系。在调查的委托人认识到

他们花掉的钱需要质量更高的回报之前,那些未能充分发挥电话调查可以对数据搜集过程进行严格控制的优点的调查,可能还会继续进行(虽然这可能是无意的)。1984 年,本书第 1 版面世之时,我就表达过这样的观点,令人遗憾的是,这种情况至今并未发生什么改变。

本书想向读者介绍的是这样一些方法,它们假定调查研究者希望在合理的费用下,搜集到高质量的数据——研究结论中的总调查误差相对较低。诚如大家所知,减少电话调查的费用,在大多数情况下都意味着降低调查质量,而降低质量就可能增加总调查误差。然而低质量的调查和不恰当的督导实际上比高质量的调查费用更高,尤其是在调查质量低到数据完全无法使用时,情况更是如此。

总之,把金钱和精力花在低质量的调查数据的搜集纯粹是一种浪费。相反,把资源用于支持精心设计和认真执行的能搜集到高质量数据的调查,它所能得到的回报往往数以倍计。

练　习

1)设计一张用于在电话上选拔调查员的表格。

2)为下面的调查计算每周应该安排的调查单元数和调查应该聘用的调查员人数:调查必须在 3 周内完成;最终完成的问卷数为 1 500 份;每个调查员平均每小时可完成问卷 1.5 份;每个调查单元为 3 小时,在集中化的访谈工作室装有 8 部电话;调查员每周平均工作 4 个工作单元。

3)为新聘的督导员撰写一份访谈单元督导员工作职责,字数不超过 500 字。

4)设计一个长度为一页的调查员监控表(IMF),对调查员使用封闭式和开放式题项的情况进行评估。

5)设计一张可以用于复查完成的问卷的表格,表格应包括一段用于解释再次与被调查人联系的原因的引导语。

术语表

Glossary of Terms

入户线路（*Access line*）：每一抽样单位（即住户）中的一条带有电话设备，且单独带有一个电话号码的电话线路。根据这一定义，那些没有电话设备的住户可视为无入户线路，而那些有一个以上电话号码的住户，则可视为有多条入户线路。

加数法（*Added-digit(s)*）：一种用于产生抽样库的技术。这种技术通常都先从一本含人名和地址的电话号码簿中选取抽样种子数，然后再在每一种子数上加上一个定数或随机数，以形成一个进入抽样库的数。例如，869-5025 是一个从某一电话簿中选出的作为种子的数，而定数 1 则是加到每一种子之上的数，那么进入抽样库的数便是 869-5026。

事先联系（*Advance contact*）：在拨打电话进行调查之前，先通过邮件或传真与抽到的被调查人或住户联系。事先联系的目的在于让被调查人对即将进行的电话调查作“预热”，以期待他们合作，从而降低无回答率。

贝尔中心（*Bellcore*）：美国电话电报公司（AT&T）所属的一个单位（1992 年该单位位于新泽西州的茅瑞斯顿（Morristown））。该单位定期编辑一种有全美地区代码和电话前置数码信息的计算机化的数据库。1992 年这种数据库以 400 美元的价格公开出售，且含有 50 000 条以上的记录。

生日选择法（*Birthday selection methods*）：一种无强制性的（nonintrusive）被调查人选择法。该法在每一抽样单位选择一

位被调查人时，以单位中所有合乎条件的人员中最近的一个，或下一个要过生日的人作为被调查人。在理论上讲，这样的选择方法应当是一种随机的选择法，但实际上这样的方法似乎并不能产生一个真正的随机样本。

回呼（*Callback*）：在与上一次呼叫的不同日子的相同或不同时间，重拨同一个电话号码。重拨的目的在于减少因为没有联系到被调查人而产生的无回答误差的可能性。

来电显示（*Caller ID*）：一种在电话通讯中可选择的服务项目。在那些视这一服务项目为合法的州中，电话用户可以预订这一选项。凡预订了这一选项的电话，每当有来电时，来电的号码便会在电话机上显示出来。这样接听者便可看到来电的号码。不难想象，这一服务项目可被某些人用来筛选掉那些他们不愿接听的电话。

来电转移（*Call forwarding*）：一种少数电话用户预订的电话通讯软件服务。这种服务允许将来电通过软件自动转移到另外一个电话号码上，因而接听来电的将是转移过去的号码，而非原来的号码。来电转移有可能使电话调查的抽样概率发生畸变。

来电筛选（*Call screening*）：无论使用来电显示还是应答机，我们都可以不接电话就知道打电话的人是谁。如果接电话的人不想接听打来的电话，那么他或她便可将那一来电筛选掉。

呼叫单（*Call-sheet*）：单独一页印有电话号码的纸。纸上的每一个电话号码都来自纸笔访谈式的电话调查的抽样库。在没有完整的计算机辅助电话访谈系统时，呼叫单为我们提供了一种用纸和笔控制抽样库的方法。通常呼叫单上记录的主要信息包括：日期、时间和每次拨号的具体处理结果，此外它还包括调查员的身份编号和其他有关的说明。有些调查组织将这些呼叫单称之为调查员报告表（*interviewer report forms*）。

计算机辅助电话访谈（*CATI*）：是指那些访谈和/或抽样库的处理的具体操作过程由计算机终端自动完成，而非由调查员以纸笔方式完成的电话调查。这些调查系统通过事先编好的程序，令计算机完成电话调查中各个阶段的工作。这些工作包括：用计算机产生或控制抽样库；提示调查员在访谈中使用适当

的语言和被调查人选择步骤;根据前提的答案进行适当的跳答;逐项显示问卷中的问题;将被调查人给出的答案以计算机可以识别的格式直接写入数据文件等。

冷呼(*Cold-call*):事先未曾通知(邮件或电传)可能的被调查人,他/她已被选为电话调查的被调查人,而直接给他们打电话进行调查。

抽样库控制者(*Controller of the sampling pool*):负责从抽样库中分发电话号码,并最终决定如何处理这些号码的人。

拒访转变(*Conversion of refusals*):为了降低无回答的数量,再次或多次重拨最初拒绝接受访问的那些被调查人,以减少拒访者的人数。其目的在于提高最终样本的外在效度(代表性)。在这一工作中,拒访报告表记录的信息将为我们提供帮助。

范围误差(*Coverage error*):总调查误差的一部分。它是因为没有给总体的每一元素提供非零的选择概率而引起的。范围误差是总体中未给出被抽取概率的那部分元素在总体中所占比例大小的函数。这一比例不同于被抽出的样本在总体中所占的比例。

拨号处理(*Dialing disposition*):对任何一次来自抽样库的电话号码的拨号尝试的处理结果。处理结果通常都用数字代码系统地记录在呼叫单上。

处理结果码(*Disposition code*):被指定用来表示某种特定的拨号结果的数字(如通了但没人接电话、废号、拒绝回答等),通常我们用这些特定的拨号结果来管理抽样库。

外在效度(*External validity*):我们可以将在样本中得到的结果推广到总体,称为该样本所代表的总体的精确度。这一精确度取决于诸多因素,如抽样设计、范围误差、无回答误差和抽样误差等。

回撤陈述(*Fallback statements*):在调查员被住户接电话的人或被调查人问及有关调查的、调查研究者根据以往经验事先已经想到的一些问题(如"你是从哪里得到我的电话号码?")时,为他们提供的标准答案。

快速忙音(*Fast-busy*):一种在拨号之后有时会听到的与忙音显然

不同的重复信号,它比一般的忙音的速度明显要快,这种快速忙音通常表示被拨的号码目前不在运行之中,或该次拨号因线路过于繁忙没有完全被执行。

看门人(*Gatekeeper*):任何一个调查员必须“通过”他/她,才能与选中的被调查人交谈的人。在通常情况下,在看门人并非调查选中的被调查人时,他/她都是第一个接电话的人。

地理位置筛选(*Geographic screening*):是简介/选择单上列出的一系列问题中的一部分,旨在确定指定的被调查人是否居住在电话调查抽样区域的地理边界内。地理位置筛选容易发生遗漏误差(伪负)和记录错误(伪正),因而使用不当的情况时有发生。

击中率(*Hit-rate*):抽样库中产生合乎条件的被调查人的电话号码的比例。在一个一般的总体调查中,击中率通常就是抽样库可涵盖的住户的百分比。

住户报告人(*Household informant*):住户中任何一个可以对调查员提出的有关住户的问题提供有效答案的人。例如任何一个可以精确地告诉调查员在住户中生活的确切人数,且该问题又是挑选出来要求被调查人回答的一系列问题中的一部分的人,都可作为住户报告人。

调查员监察表(*Interviewer monitoring form*, *IMF*):一种督导人员使用的对调查进行监察的表格。记录在 IMF 上的信息既可反馈给调查员,也可被研究人员用来评估与调查员有关的测量误差的大小和性质。

调查员生产率(*Interviewer productivity*):在一个给定的时间范围内(例如 1 小时)调查员按要求完成的访谈人数和在同一时间内拒访人数之比。一般来讲,完成/拒绝比不能低于 3:1,调查员必须为达到这一目标而努力。

简介/选择单(*Introduction/selection sheet*):一种独立的纸质清单,常附于呼叫单之后,内容多为一些引导性用语和在调查中应用的选择被调查人的技巧。

引导用语(*Introductory spiel*):在与一个可能的符合条件的住户或被调查人联系时,由调查员宣读的一种标准化的引导性用语。

凯思选择法（*Kish selection method*）：一种最为严格的被调查人选择法。凯思法是被人们普遍接受的用于个人调查的调查对象的选择方法。它要求列出抽样单位内所有符合调查标准（如18岁或18岁以上）的个人。虽然这是一种在理论上作了精心考虑的真正的概率选择法，但是它仍然无法保证我们在实际的电话调查中得到真正随机的单位样本。在调查员开始与调查单位进行接触时，这种方法会令人感到被过多地打扰而拒绝接受访谈。那些缺乏经验的调查员往往因为使用这种方法而导致非常高的拒访率。

误差差额（*Margin of error*）：见抽样误差。

测量误差（*Measurement error*）：总调查误差的一部分。由调查员或被调查人的行为及问卷和/或调查实施的方式引起。从某种程度上讲，测量误差是无法完全排除的。如果可能，一个行事谨慎的研究者应当在调查中构建一种方法论意义上的检验，使我们可以在调查完成后了解测量误差的大小。

米托夫斯基-瓦克舍尔二阶随机数码拨号抽样法（*Mitofsky-Waksberg two-stage RDD sampling*）：一种为人们所普遍接受的抽样方法。它可以产生一个非常高的住户击中率的随机数码拨号（RDD）抽样库。在做一阶抽样时，我们用简单随机数码拨号法，确定相对较少的进入抽样区内的住户的电话号码。这些号码的前置号和它们后置号中的前几位，通常是前两位一起作为产生二阶抽样库的种子。与一阶抽样相比，二阶抽样的抽样误差会稍高一些。

混合式付酬（*Mixed mode payment*）：一种调查员的酬金由固定计时工资和随生产效率变化的奖金两部分组成的付酬方式。

混合式调查（*Mixed mode survey*）：使用一种以上调查方式（邮寄调查、个别访谈和电话调查）为同一调查项目搜集数据。采用混合式调查的目的在于取各种调查方式之长，以在尽可能减少各种调查方式固有缺点的同时减少总调查误差。

多框抽样（*Multiple-frame sampling*）：在一个调查中，抽样库的产生使用了一个以上的抽样框。例如，将随机数码拨号抽样和从反向电话簿抽样两种方法结合在一起进行的抽样。试图在调

查费用和可能的误差之间找到一个理想的平衡点。

间接探问(*Nondirective probing*):对模糊不清的答案作进一步的探问,否则将会因为调查员未予深究而得到不够精确的答案。这种提问技巧主要关注的是,被调查人在给出问题的答案时所表现出来的行为方式,而不是他们在这些答案中表达出来的思想认识。间接探问的一个例子是:"您是不是可以把最后一点再说得确切一点?"

无回答误差(*Nonresponse error*):总调查误差的一部分。它是一种由那些已经抽到但没有搜集到数据的样本单元,与那些不仅抽到且搜集到数据的样本单元之间存在的差异而引起的误差。调查中无回答问题产生的主要原因是,无法同被调查人取得联系和被调查人拒绝回答,其中后者构成了公众电话调查中的大多数。无回答误差大小是抽到但没有调查到的那部分人在总体所占比例大小,和这部分人与不仅抽到而且调查到的那部分人之间差异大小的函数。如果"无回答者"就调查所搜集的测量指标而言,与"回答者"并不存在什么差异,那么无回答误差便不复存在。

非抽样误差(*Nonsampling error*):一种与范围误差、测量误差和无回答误差有关的复合误差。

在岗培训(*On-the-job training*):在调查阶段,督导员通过结构性反馈对调查员进行的培训。

原样本追踪调查(*Panel survey*):一种对相同的调查对象进行再次或多次调查的调查方式。每轮调查使用基本相同的问卷,两轮的间隔通常为 12 个月。

纸笔访谈(*PAPI*):PAPI 是英文纸笔访谈(Paper-and-pencil-interviewing)的缩写。访谈中调查员拿一份问卷,逐一宣读印在上面的问题,并将问题答案直接记录在问卷上。

探索试验(*Pilot test*):一种数量相对较少的实际访谈,目的在于试验并进一步改进抽样库、引导性用语、选择步骤和整个问卷。探索试验也可帮助我们确定完成一份问卷所需的时间。通常,在完成探索检验之后,应举行一个专门的会议来听取那些参加实际试验的调查员的情况汇报。

总体参数（*Population parameter*）：某种代表某一变量在总体中存在水平的值。我们常常用调查数据来估计这些参数值。

前置数字（*Prefix*）：地方电话号码的前三位数字，通常用来代表某地的电话交换站。

概率样本（*Probability sample*）：用各种概率抽样得到的样本。所谓概率抽样是指那些能在理论上给总体中的每一元素提供一个已知的非零的选择概率的抽样方法。只有概率样本才可以在数量上计算抽样误差的大小。

质量控制（*Quality control*）：一种在调查中用来提高抽样和访谈质量的系统方法，借助这种方法，我们可以减少总调查误差。一般来讲，电话调查较之其他形式的调查为我们提供了更多的质量控制的可能。

随机数字拨号（*Random digit dialing*）：若干种在呼叫的抽样区域的前置号码已知的情况下，通过在前置号后面加入随机数字来生成抽样库的方法。我们可借助这种方法，通过电话对居住在没有电话簿的地区和/或电话号码没有刊印在电话簿上的那些住户进行调查。

随机数字表（*Random numbers table*）：在许多统计书籍后面刊出的印有数千个数字(0～9)的表。这些数字以某种随机的次序刊印。这些随机数字串对调查研究者可有多种用途，例如，可用它们为前置数生成用于随机数字号码抽样库中的随机后置数。

避免拒访培训（*Refusal avoidance training*）：一种培训。目的是使调查员理解和掌握如何在实际调查过程中避免各种原因的拒访的发生。培训的主要内容是向调查员介绍各种与降低拒访率有关的策略。模拟实际调查也是这种培训的重要内容之一。

拒答报告表（*Refusal report form*，*RRF*）：在调查员每次遇到拒答时都需填写的结构性表格。这种表格有助于随后进行的拒答转变尝试，不仅如此，它也有助于研究者了解调查中出现的巨大误差的性质。

被调查人排除标准（*Respondent exclusion criterion*）：某些事先确定的用于排除不合格的被调查人如年纪太小、或不能说英语等的

标准。

被调查人选择法(Respondent selection):一种调查员在开场白之后用来在一个抽样单元的所有合格被调查人中挑选一个合适的被调查人的技巧。

回答率(Response rates):若干种计算效率的数量量度(百分比)的方法。这些方法包括抽样库在达到调查规定的合格的被调查人人数时的效率计算方法,以及调查员在完成这些合格的被调查人的访谈时的效率计算方法等。

反向电话簿(Reverse directory):一种特殊的以数字顺序和/或以住址排列,而非以姓氏的字母次序排列的电话簿。反向电话簿不含没有登记的电话。

角色扮演(Role playing):一种培训调查员的技巧。在培训中要求调查员模拟进行各种实际操作,如陈述调查简介、进行户内被调查人选择和逐项宣读问卷题项等。被调查人的角色一般都由督导员担当。

抽样边界(Sampling boundaries):一个特定的地理区域,电话调查的单位或被调查人将在其中抽取。

抽样设计(Sampling design):指导我们从研究总体中抽取调查样本的计划。它需要确定的内容有:目标总体、抽样边界、如何从抽样框中产生抽样库、如何从单元中选择回答人、最终样本的大小和抽样后的加权修正。

抽样误差(Sampling error):总调查误差的一部分。与普查不同,抽样调查研究的只是总体所有元素的一部分,即样本,抽样误差便因此而产生。如果我们使用的样本是一个概率样本,那么它便可以用标准的统计公式计算出来。

抽样框(Sampling frame):抽样库赖以产生的登录目标总体的各种元素的清单。在随机数码拨号调查中,抽样框只在理论上存在。从某种意义上讲,它是由数以万计的、所有能在抽样区域内接通的前置号的每一前置号和它所有可能存在的后置号(即0000~9999)组成的所有电话号码构成的。

抽样库(Sampling pool):一组实际的电话号码,产生于抽样框中,调查员用它与被调查人进行联系。在纸笔式调查中,使用的来

自抽样库的每一个电话号码,都刊印在一张单独的呼叫单上,以防止调查员使用不当。

饱和度(*Saturation*):在将电话调查应用于一般总体时,抽样区内拥有至少一条电话线的住户的比例。20 世纪 90 年代,美国居民的电话饱和度估计在 93% ~95% 。

跳跃模式(*Skip pattern*):问卷题项的一种排列模式。在这种排列模式中,后面应该回答的题项的选择,取决于被调查人在前面的某些题项中给出的一个或若干个答案。

前置数分层(*Stratification by prefix*):一种抽样库。在这种抽样库中,不同前置数的电话码数的比例,与它们在抽样边界内的比例相一致。

后置数字(*Suffix*):电话号码中的最后四位数字。

骗销术(*Sugging*):一种不道德的电话销售技巧,即所谓的"伪装式销售"。在这样一种电话销售中,尽管呼叫者的目的始终是推销商品,但他们却把自己装扮成电话调查的调查员。有人认为骗销术的流行导致了一般公众拒答问题的比例上升。

系统抽样(*Systematic sampling*):不同于直接从电话簿或其他清单中形成抽样库的方法。这种方法首先要确定一个随机的起始点,然后再以固定的间距,自起始点开始,逐个从清单上选出进入抽样库的下一个单元。

目标总体(*Target population*):样本实际抽取的有限总体。

总调查误差(*Total survey error*):调查中发现存在的偏倚(恒定误差)和变差(不定误差)。它们由范围误差、无回答误差、测量误差和抽样误差的影响综合所致。从某种意义上讲,总调查误差是不可避免的,但它的大小和性质是可以测量的。

未登录号码(*Unlisted numbers*):任何既未收录在正式出版的电话簿,又未收录在电话号码查询资料上的电话号码,都被称之为未登录号码。

未出版号码(*Unpublished numbers*):任何未收录在正式出版的电话簿上的电话号码。它是否收录在电话号码查询资料上,取决于当地电话公司的管理程序。

空置后置数字段(*Vacuous suffix bank*):某一给定的前置号码的未曾

使用的后置数字段。

审查(*Validation*):在纸笔式电话调查中,值班的督导员在当班的访谈单元中,对已经完成的问卷进行的一种即时审查。审查的内容包括整个问卷的完成情况是否恰当,以及开放性问题答案的记录字迹是否清楚等。

复查(*Verification*):再次致电被调查人,以确认调查员是否确实进行过访谈。复查由督导人员操作完成,一般在原调查完成之后一两天内进行。

可视电话(*Videophone*):一种从 1992 年以后开始向公众出售的电话电讯产品。这种产品运行在一条常备线路上,通话双方都有一部可视电话,在通话的同时,彼此可以看见对方。如果这种技术为公众所普遍接受,那么电话调查工作者就必须采用和发展一些新的方法,以满足使用这些可减少总误差的技术的需要。

加权(*Weighting*):一种在事后对调查数据进行修正的统计方法,目的是增加数据的外在效度,进而增加对总体参数所作估计的精确性。在电话调查中,对由多入户线路和抽样单位内合格的被调查对象人数的不确定性引起的选择不等概问题进行加权不仅合乎规范,而且也是切实可行的。为了对无回答问题作一些补救,我们经常对最终样本和总体之间存在的人口学上的差别进行加权(但这种做法未必总是恰当的)。

单元内抽样(*Within-unit sampling*):结构性被调查人选择法。主要用于在抽样单元(如住户)内选择特定数目的(通常是一个)合乎要求的被调查人。

参考文献

References

American Association for Public Opinion Research (AAPOR). (1991). *Code of professional ethics and practices.* Ann Arbor, MI: AAPOR Secretary's Office.

Andrews, E. L. (1992, March 15). Emboldened phone companies are pushing the frills. *New York Times*, p. F8.

Alexander, C. A. (1988). Cutoff rules for secondary calling in a random digit dialing survey. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 113-126). New York: John Wiley.

Alexander, C. A., Sebold, J., & Pfaff, P. (1986). Some results with an experiment with telephone sampling for the US national crime survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, Chicago.

AT&T. (1982). *The world's telephones*: 1982. Atlanta: R. H. Donnelley.

Babbie, E. (1989). *The practice of social research* (5th ed.). Belmont, CA: Wadsworth.

Baker, R. P., & Lefes, W. L. (1988). The design of CATI systems: A review of current practice. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 387-402). New York: John Wiley.

Bass, R. T., & Tortora, R. D. (1988). A comparison of centralized CATI facilities for an agricultural labor survey. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 497-508). New York: John Wiley.

Bauman, S. L., Merkle, D. M., & Lavrakas, P. J. (1992). *Interviewer estimates of refusers' gender, age, and race in telephone surveys.* Paper presented at the Midwest Association for Public Opinion Research Conference, Chicago.

Belson, W. A. (1981). *The design and understanding of survey questions.* Aldershot, England: Gower.

Berry, S. H., & O'Rourke, D. (1988). Administrative designs for centralized telephone survey

centers: Implications of the transition to CATI. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 457-474). New York: John Wiley.

Biemer, P. P., Groves, R. M., Lyberg, L. E., Mathiowetz, N. A., & Sudman, S. (Eds.). (1991). *Measurement errors in surveys.* New York: John Wiley.

Blair J., & Chun, Y. (1992). *Quality of data from converted refusals in telephone surveys.* Paper presented at the American Association for Public Opinion Research Conference, St. Petersburg, FL.

Bradburn, N. M., & Sudman, S. (1991). The current status of questionnaire design. In P. P. Biemer. R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 29-40). New York: John Wiley.

Bradburn, N. M., Sudman, S., & Associates. (1979). *Improving interview method and questionnaire design.* San Francisco: Jossey-Bass.

Brook, J. (1976). *Telephone: The first hundred years.* New York: Harper & Row.

Bryant, B. E. (1975). Respondent selection in a time of changing household composition. *Journal of Marketing Research*, *12*, 129-135.

Burkheimer, G. J., & Levinsohn, J. R. (1988). Implementing the Mitofsky-Waksberg sampling design with accelerated sequential replacement. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 99-112). New York: John Wiley.

Catlin, G., & Ingram, S. (1988). The effects of CATI on costs and data quality: A comparison of CATI and paper methods in centralized interviewing. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 437-452). New York: John Wiley.

Campbell, D. T., & Stanley, J. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Cannell, C. (1991, October). *Research on monitoring interviewers' behaviors.* Informal presentation at the Institute for Social Research, University of Michigan.

Cannell, C., & Oksenberg, L. (1988). Observation of behavior in telephone interviews. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 475-496). New York: John Wiley.

Cochran, W. G. (1977). *Sampling techniques.* New York: John Wiley.

Collins, M., Sykes, W., Wilson, P., & Blackshaw, N. (1988). Nonresponse: The UK experience. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 213-232). New York: John Wiley.

Congressional Information Service (1990). *American statistical index.* Bethesda, MD: CIS.

Cook, T. D., &Campbell, D. T. (1979). *Quasi-experimentation: Designs and analysis issues for field settings.* Chicago: Rand McNally.

Cooper, S. L. (1964). Random sampling by telephone: An improved method. *Journal of Marketing Research*, *1*(4), 45-48.

Couper. M. P., Holland, L., & Groves, R. M. (1992). "*Hello my name is ...*": *Respondentinterviewer interactions in survey introductions*. Paper presented at the American Association for Public Opinion Research Conference, St. Petersburg, FL.

Czaja, R., Blair, J., & Sebestik, J. (1982). Respondent selection in a telephone survey. *Journal of Marketing Research*, *19*, 381-385.

Dawes, R. M. (1972). *Fundamentals of attitude measurement*. New York: John Wiley. de Leeuw, E. D., & van der Zouwen, J. (1988). Data quality in telephone and face to face surveys: A comparative meta-analysis. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 283-300). New York: John Wiley.

de Sola Pool, I. (Ed.). (1977). *The social impact of the telephone*. Cambridge: MIT Press.

Dillman, D. A. (1978). *Mail and telephone surveys: The total design method*. New York: John Wiley.

Dillman, D. A., Gallegos, J., & Frey, J. H. (1976). Reducing refusals for telephone interviews. *Public Opinion Quarterly*, *40*, 99-114.

Dillman, D. A,, Sangster, R. L., & Rockwood. T. H. (1992). *Question form effects in mail and telephone surveys: Results from 14 experiments*. Paper presented at the American Association for Public Opinion Research Conference, St. Petersburg, FL.

Dillman, D. A,, & Tarnai, J. (1988). Administrative issues in mixed mode surveys. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey. W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 509-528). New York: John Wiley.

Dillman, D. A,, & Tarnai, J. (1991). Mode effects of cognitively designed recall questions: A comparison of answers to telephone and mail surveys. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 73-94). New York: John Wiley.

Drew, J. D., Choudhry, G. H., & Hunter, L. A. (1988). Nonresponse issues in government telephone surveys. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephonesurvey methodology* (pp. 233-246), New York: John Wiley.

Ekman, P., & Friesen, W. (1974). Detecting deception from body and face. *Journal of Personality and Social Psychology*, *29*, 288-298.

Ekman, P., & Friesen, W. (1976). Body movement and voice pitch in deceptive interaction. *Semiotica*. *16*, 23-27.

Fowler, F. J., Jr. (1984). *Survey research methods*. Beverly Hills, CA: Sage.

Fowler, F. J., Jr. (1993). *Survey research methods* (2nd ed.). Newbury Park, CA: Sage.

Fowler, F. J., Jr., & Mangione, T. W. (1985). *The value of interviewer training and*

supervision. Final Report to the National Center for Health Services Research, Washington, DC.

Fowler, F. J., Jr., & Mangione, T. W. (1986). *Reducing interviewer effects on health survey data.* Washington, DC: National Center for Health Statistics.

Fowler, F. J., Jr., & Mangione, T. W. (1990). *Standardized survey interviewing.* Newbury Park, CA: Sage.

Frey, J. H. (1989). *Survey research by telephone* (2nd ed.). Newbury Park, CA: Sage.

Gawiser, S. R., & Witt, G. E. (1992). *Twenty questions a journalist should ask about poll results.* New York: National Council on Public Polls.

Groves, R. M. (1989). *Survey errors and survey costs.* New York: John Wiley.

Groves, R. M., Biemer, P. P., Lyberg, L. E., Massey, J. T., Nicholls, W. L., & Waksberg, J. (Eds.). (1988). *Telephone survey methodology.* New York: John Wiley.

Groves, R. M., & Kahn, R. L. (1979). *Surveys by telephone: A national comparison with personal interviews.* New York: Academic Press.

Groves, R. M., & Lyberg, L. E. (1988). An overview of nonresponse issues in telephone surveys. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 191-212). New York: John Wiley.

Hagen, D. E., & Collier, C. M. (1982). Must respondent selection procedures for telephone surveys be invasive? *Public Opinion Quarterly*, *47*, 547-556.

Hays, W. L. (1973). *Statistics for the social sciences.* New York: Holt, Rinehart & Winston.

Hedrick, T., Bickman, L., & Rog, D. (1993). *Applied research design.* Newbury Park, CA: Sage.

Henry, G. T. (1990). *Practical sampling.* Newbury Park, CA: Sage.

House, C. C., & Nicholls, W. L. (1988). Questionnaire design for CATI: Design objectives and methods. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 421-436). New York: John Wiley.

Kish, L. (1949). Aprocedure for objective respondent selection within the household. *Journal of the American Statistical Association*, *44*, 380-387.

Kish, L. (1962). Studies of interviewer variance for attitudinal variables. *Journal of the American Statistical Association*, *57*, 92-115.

Kish, L. (1965). *Survey sampling.* New York: John Wiley.

Kviz, F. J. (1977). Towards a standard definition of response rate. *Public Opinion Quarterly*, *41*, 265-267.

Landon, E. L., & Banks, S. K. (1977). Relative efficiency and bias of plus-one telephone sampling. *Journal of Marketing Research*, *14*, 294-299.

Lavrakas, P. J. (1987). *Telephone survey methods: Sampling, selection, and supervision.* Newbury Park, CA: Sage.

Lavrakas, P. J. (1990). *Morton Grove district 70 enrollment study* [Mimeo]. Evanston, IL: Northwestern University Survey Lab.

Lavrakas, P. J. (1991). Implementing CATI at the Northwestern Survey Lab: Part I. *CATI News*, 4(1), 2-3 +.

Lavrakas, P. J. (1992). *Attitudes towards and experiences with sexual harassment in the workplace.* Paper presented at the Midwest Association for Public Opinion Research Conference, Chicago.

Lavrakas, P. J., & Maier, R. A. (1979). Differences in human ability to judge veracity from the audio medium. *Journal of Research in Personality, 13*, 139-153.

Lavrakas, P. J., & Merkle, D. M. (1991). *A reversal of roles: When respondents question interviewers.* Paper presented at the Midwest Association for Public Opinion Research Conference, Chicago.

Lavrakas, P. J., Merkle, D. M., & Bauman, S. L. (1992). *Refusal report forms, refusal conversions, and nonresponse bias.* Paper presented at the American Association for Public Opinion Research Conference, St. Petersburg, FL.

Lavrakas, P. J., Merkle, D. M., & Bauman, S. L. (1993). *The last-birthday selection method and within-unit coverage problems.* Paper presented at the American Association for Public Opinion Research Conference, St. Charles, IL.

Lavrakas, P. J., & Rosenbaum, D. P. (1989). *Crime prevention beliefs, policies and practices of chief law enforcement executives: Results of national study.* Washington, DC: National Crime Prevention Council.

Lavrakas, P. J., Settersten, R. A., Jr., & Maier, R. A., Jr. (1991). RDD panel attrition in two local area surveys. *Survey Methodology, 17*, 143-152.

Lavrakas, P. J., Skogan, W. G., Normoyle, J., Herz, E. J., Salem, G., & Lewis, D. A. (1980). *Factors related to citizen participation in personal, household, and neighborhood anticrime measures* [Mimeo]. Evanston, IL: Center for Urban Affairs & Policy Research.

Lavrakas, P. J., & Tyler, T. R. (1983). *Low cost telephone surveys.* Paper presented at Evaluation '83, Chicago.

LeBailly, R. F., & Lavrakas, P. J. (1981). *Generating a random digit dialing sample for telephone surveys.* Paper presented at Issue '81 Annual SPSS Convention, San Francisco.

Lepkowski, J. M. (1988). Telephone sampling methods in the United States. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 73-98). New York: John Wiley.

Lewis, I. A. (1991). Media polls, the *Los Angeles Times* poll, and the 1988 presidential election. In P. J. Lavrakas & J. K. Holley (Eds.), *Polling and presidential election coverage* (pp. 57-82). Newbury Park, CA: Sage.

Lipsey, M. W. (1989). *Design sensitivity: Statistical power for experimental research.* Newbury Park, CA: Sage.

Lyberg, L. E. (1988). Introduction: The administration of telephone surveys. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 453-456). New York: John Wiley.

Lyberg, L. E., & Dean, P. (1992). *Methods for reducing nonresponse rates—a review.* Paper presented at the American Association for Public Opinion Research Conference, St. Petersburg, FL.

Maier, N. R. F. (1966). Sensitivity to attempts at deception in an interview situation. *Personnel Psychology*, *19*, 55-66.

Maier, N. R. F., & Thurber, J. (1968). Accuracy of judgments of deception when an interview is watched. heard or read. *Personnel Psychology*, *21*, 23-30.

Maier, R. A., & Lavrakas, P. J. (1976). Lying behavior and evaluation of lies. *Perceptual and Motor Skills*, *42*, 575-581.

Maklan, D., & Waksberg, J. (1988). Within-household coverage in RDD surveys. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 51-72). New York: John Wiley.

Mason, R. E., & Immerman, F. W. (1988). Minimum cost sample allocation for Mitofsky-Waksberg random digit dialing. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 127-142). New York: John Wiley.

Massey, J. T. (1988). An overview of telephone coverage. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 3-8). New York: John Wiley.

Massey, J. T., & Botman, S. L. (1988). Weighting adjustments for random digit dialed surveys. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 143-160). New York: John Wiley.

Merkle, D. M., Bauman, S. L., & Lavrakas, P. J. (1991). *Nonresponse bias: Refusal conversions and call-backs in RDD telephone surveys.* Paper presented at the Midwest Association for Public Opinion Research Conference, Chicago.

Mitofsky, W. J. (1970). *Sampling of telephone households* [Mimeo]. New York: CBS News.

Nicholls, W. L. (1988). Computer-assisted telephone interviewing: A general introduction. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone Survey Methodology* (pp. 377-386). New York: John Wiley.

Oldendick, R. W., Sorenson, S. B., Tuchfarber, A. J., & Bishop, G. F. (1985). *Last birthday respondent selection in telephone surveys: A further test.* Paper presented at Midwest Association of Public Opinion Research, Chicago.

Oksenberg, L., & Cannell, C. (1988). Effects of interviewer vocal characteristics on nonresponse. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 257-272). New York: John

Wiley.

O'Rourke, D., & Blair, J. (1983). Improving random respondent selection in telephone surveys. *Journal of Marketing Research*, *20*, 428-432.

Pierce, J. R. (1977). The telephone and society in the past 100 years. In I. de Sola Pool (Ed.), *The social impact of the telephone* (pp. 159-195). Cambridge: MIT Press.

Potthoff, R. F. (1987). Some generalizations of the Mitofsky-Waksberg technique for random digit dialing. *Journal of the American Statistical Association*, *82*, 409-418.

Ramirez, A. (1992, April 30). Lifetime telephone numbers that ring anywhere you go. *New York Times*, p. Al +.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. W. (1991). *Measures of personaliv and social psychological attitudes.* New York: Academic Press.

Salmon, C. T., & Nichols, J. S. (1983). The next-birthday method for respondent selection. *Public Opiniom Quarterly*, *47*, 270-276.

Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys.* New York: Academic Press.

Sebold, J. (1988). Survey period length, unanswered numbers, and nonresponse in telephone surveys. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 247-256). New York: John Wiley.

Smead, R. J., & Wilcox, J. (1980). Ring policy in telephone surveys. *Public Opinion Quarterly*, *44*, 115-116.

Stokes, L., & Ming-Yih, Y. (1988). Searching for causes of interviewer effects in telephone surveys. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 357-376). New York: John Wiley.

Sudman, S. (1973). The uses of telephone directories for survey sampling. *Journal of Marketing Research*, 10 (May), 204-207.

Sudman, S. (1976). *Applied sampling.* New York: Academic Press.

Sudman, S., & Bradburn, N. (1974). *Response effects in surveys: A review and synthesis.* Chicago: Aldine.

Sudman, S., & Bradburn, N. (1982). *Asking questions.* San Francisco: Jossey-Bass.

Survey Sampling, Inc. (1990). *A survey researcher's view of the U. S.* Fairfield, CT: Survey Sampling.

Thornberry, O. T., Jr., & Massey, J. T. (1988). Trends in United States telephone coverage across time and subgroups. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey mefhodology* (pp. 25-50). New York: John Wiley.

Traugott, M. R. (1987). The importance of persistence in respondent selection for preelection surveys. *Public Opinion Quarterly*, *51*, 48-57.

Traugott, M. R., Groves, R. M., & Lepkowski, J. (1987). Using dual frame designs to reduce

nonresponse in telephone surveys. *Public Opinion Quarterly*, *51*, 522-539.

Trewin, D., & Lee, G. (1988). International comparisons of telephone coverage. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 9-24). New York: John Wiley.

Troldahl, V. C., & Carter, R. E., Jr. (1964). Random selection of respondents within households in phone surveys. *Journal of Marketing Research*, *1*, 71-76.

Tucker, C. (1983). Interviewer effects in telephone surveys. *Public Opinion Quarterly*, *47*, 84-95.

U. S. Bureau of Census (1984). *Statistical abstracts of the United States*, 1984. Washington, DC: Government Printing Office.

Waksberg, J. (1978). Sampling methods for random digit dialing. *Journal of the American Statistical Association*, *73*, 40-46.

Weeks, M. F. (1988). Call scheduling with CATI: Current capabilities and methods. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls. & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 403-420). New York: John Wiley.

Weeks, M. F., Jones, B. L., Folsom, R. E., & Benrud, C. H. (1980). Optimal times to contact sample households. *Public Opinion Quarterly*, *44*, 101-114.

Weeks, M. F., Kulka, R. A., & Pierson, S. (1987). Optimal call scheduling for a telephone survey. *Public Opinion Quarterly*, *51*, 540-549.

Weiss, C. (1972). *Evaluation research.* Englewood Cliffs, NJ: Prentice-Hall.

Wolter. K. (1985). *Introduction to variance esttmation.* New York: Springer Verlag.